HANFORD THYROID DISEASE STUDY
FINAL REPORT

Study Management Team

Scott Davis, Ph.D., Principal Investigator (Epidemiology)
Kenneth J. Kopecky, Ph.D., Co-Investigator (Biostatistics)
Thomas E. Hamilton, M.D., Ph.D., Co-Investigator (Endocrinology)

This Report Was Prepared by the HTDS Study Management Team
And the Following HTDS Staff

Lynn E. Onstad, Sc.M., Statistical Research Associate
Beth L. King, B.A., Project Manager
Mark S. Saporito, M.A., Systems Analyst

With Special Assistance From
Christy R. Callahan, B.A.

Fred Hutchinson Cancer Research Center

CDC Contract Number 200-89-0716

June 21, 2002

Acknowledgments

Many individuals have contributed to the success of the Hanford Thyroid Disease Study. We wish to express our deep appreciation to Ms. Peggy Adams Myers for her leadership as Project Manager from the beginning of the study through the release of the Draft Final Report, and to Ms. Beth King as Project Manager thereafter. We also extend a special acknowledgment to several persons who supervised or were responsible for major components of the study: Lynn Onstad (data operations and statistical analysis); Mark Saporito (programming); Sandy Walbrek (administrative support); Beth King (tracing and recruiting); Theresa Taggart (travel); Esther Kohler (CATI); Surena Khatir (database support); and Thula Weisel and Laurie Shields (clinic operations). Ms. Dana Mirick provided invaluable statistical support.

Dr. Bruce Amundson provided expertise in family and rural medicine, and was a member of the Study Management Team through August, 1998.

We also wish to thank the many additional staff who worked on this project for their valuable contributions in the following areas: 1) administrative support: Teri Kopp, Cathy Voigt, Nancy McMillan, Sara Smith, Tricia Christopherson, Janice Findley, Iona McKenzie, Georgia Green, Kimberly Green, Berta Nicole-Blades, Carroll Haymon, Donna Johnson, Steve Walbrek, Christy Callahan; 2) tracing and recruiting: Kathy Koons, Allan Jaeger, Elizabeth Acorda, Meg Woods, Judith Dietert, Karen Queen, Dick Young, Anna Marie Biasini, Jana Carr, Carolyn Tuman, Jesse Winter; 3) CATI: Bruce Cummins, Shannon Greer, Kathy Vickers, Jaime Roberts-Jones, Eileen Bennoff, Sue Kung; 4) clinics: Steve Minor, Esther Ricardo-Bulis, Mari Hooten-Lee, Sally Kirk, Holly Roberts, Michael Weygint, Melanie Bartmess, Lynne Brosnahan; 5) thyroid ultrasound: Theresa Stanfel-Hull, David Sherman, Jeremy Hodges, Carmen Fletcher; 6) medical records: Susan DeNeef, Sherry Shanabarger; 7) data operations: Bernie Moore, Elizabeth Carosso, Meri Gilman, Jennifer Sporleder, Jan Kikuchi, Bill Mullin, Liza Noonan, Chuck Wiggins, Belen Gallardo, and especially Linda Joos (deceased).

Along with Drs. Hamilton and Griep, the following endocrinologists examined study participants in the HTDS clinics: Drs. Gayle Brewer, Ken Gross, Gary Treece, Mike Tuttle. Dr. Bruce Kulander provided pathology review and consultation for all fine needle aspiration and histology materials. Drs. Keith Wang, Jack Hirsch, John Denney, and Robert Schor provided thyroid ultrasound review and consultation for all ultrasound exams. The following consultants provided valuable assistance in the following areas: 1) CATI questionnaire design and implementation: Drs. Donald Dillman, John Tarnai, Ronald Fisher, and Ms. Ellen Lammiman; 2) birth and death certificate access and retrieval from the State of Washington: Dr. Floyd Frost and Mr. Eugene Sobota; 3) history of Hanford: Dr. Michelle Stenejhem; 4) statistical and design issues: Dr. Duncan Thomas, Dr. Dana Flanders, and Dr. Dan Stram; and 5) Anti-Tg assay and results: Dr. Carol Spencer.

We especially acknowledge the contributions of Dr. Robert W. Day, President and Director Emeritus of the Fred Hutchinson Cancer Research Center, for his support of this project and his encouragement through all stages of the research. He greatly facilitated the conduct of this complex endeavor.

We extend a special thank you to Dr. John Till, who served as the Chairman of the Technical Steering Panel (TSP) of the Hanford Environmental Dose Reconstruction Project, for his assistance and support throughout the project, and to each member of the TSP for their guidance and encouragement. Dr. David Price was particularly helpful in issues related to milk consumption and the development of the dosimetry questionnaire, and Ms. Mary Lou Blazek provided valuable assistance in matters related to public communications. Dr. Bruce Napier of Battelle Pacific Northwest Laboratories was especially helpful in developing the dosimetry (CATI) interview, and in matters related to the implementation of the computer system for estimating individual thyroid doses. We also thank the many volunteers who so kindly gave of their time to assist in the testing and revision of the interview instruments that were used for dosimetry (CATI) and in the clinic (in-person interview).

We wish to acknowledge the dedicated service and valuable contributions provided by members of the federal Advisory Committee convened to advise the CDC on the conduct of the HTDS. The following individuals served on the committee for the entire course of the study: Dr. Larry Jecha (Chairman), Dr. Arthur Schneider (thyroid specialist), Mr. Lou Stone (Native American representative). Others who contributed to the committee and who were members for varying lengths of time include: Dr. Owen Hoffman (dosimetry), Dr. Marlene McKetty (dosimetry), Dr. Genevieve Matanoski (epidemiology), Dr. Maureen Hatch (epidemiology), Ms. Kristine Gebbie (State of Washington), Ms. Elizabeth Ward (State of Washington), Mr. Jim Thomas (regional community), and Dr. Sara Cate (regional community). Two consultants representing the Hanford Downwinder's Coalition also contributed to the deliberations of the committee: Ms. Pam Hofer and Ms. Judith Jurji.

Several organizations have provided valuable assistance in the conduct of this study: Pacific Medical Center, Seattle Laboratory of Pathology, Seattle Nuclear Medicine, The Seattle Foundation, and the following sites which provided space for one or more HTDS clinics: Our Lady of Lourdes Business Health Center, Pasco, Washington; Deaconess Medical Center, Spokane, Washington; Southwest Washington Medical Center, Vancouver, Washington; Internal Medicine Associates of Yakima, Yakima, Washington; St. Mary Regional Medical Center, Walla Walla, Washington; Central Washington Hospital, Wenatchee, Washington; Mt. Carmel Hospital, Colville, Washington; Okanogan Valley Clinic, Omak, Washington; and St. Vincent Hospital and Medical Center, Portland, Oregon.

Finally, we gratefully acknowledge the substantial contributions that each and every study participant made in willingly giving of their time and personal information, as well as the contributions made by many family members. Without such unselfish dedication on the part of so many people, this study would not have been possible. The overall success of this effort is due primarily to the participants.

TABLE OF CONTENTS

LIST OF FIGURES

II. BACKGROUND

IV. STUDY DESIGN

V. FIELD PROCEDURES AND METHODS, RESULTS OF DATA COLLECTION PROCESS

VIII. STATISTICAL METHODS

IX. RESULTS

LIST OF TABLES

II. BACKGROUND

V. FIELD PROCEDURES AND METHODS, RESULTS OF DATA COLLECTION PROCESS

VIII. STATISTICAL METHODS

IX. RESULTS

Executive Summary

I. Background

The Hanford Thyroid Disease Study (HTDS) was mandated by an act of Congress in 1988. The Centers for Disease Control and Prevention (CDC) was directed by Senate Bill 2889 to conduct a study of thyroid morbidity among persons who lived near the Hanford Nuclear Site between 1944 and 1957. A team of investigators at the Fred Hutchinson Cancer Research Center (FHCRC) and the University of Washington in Seattle was selected by the CDC to conduct the study, and a contract was awarded to the FHCRC on September 19, 1989.

The primary purpose of the study was to determine whether thyroid morbidity is increased among persons exposed to releases of radioactive iodine from the Hanford Nuclear Site between 1944 and 1957. The study was also designed to further determine in what way any increase in thyroid morbidity was related to the dose of radiation received (i.e., the characteristics of any dose-response relationship). Secondary objectives of the study included the following: 1) to determine whether hyperparathyroidism is increased among persons exposed to the Hanford radiation and, if so, to determine in what way the increase is related to the dose of radiation received; 2) to provide information to residents of the communities surrounding the Hanford Site regarding the objectives, design, and conduct of the study, as well as the findings and results of the research; and 3) to assess the appropriateness of the methods employed and the degree to which such an investigation could be successfully planned and executed, given the long interval since exposure and the uncertainties regarding radiation dose.

This study was conducted as a follow-up prevalence study. That is, a group of individuals (a "cohort") was selected on the basis of presumed past exposure to varying levels of radioactive iodine (^{131}I in particular) released into the atmosphere from Hanford, based on place and year of birth. Individuals in the cohort were located and those who participated had a dose estimate calculated from answers to a dosimetry questionnaire, and were examined for the presence or history of thyroid disease. The primary analyses focused on living participants who received medical examinations to detect thyroid disease, and for whom thyroid radiation doses were estimated using the dosimetry system developed by the Hanford Environmental Dose Reconstruction (HEDR) Project. All forms of thyroid disease were investigated as part of the study and were included in the analysis, as were abnormalities of the thyroid gland seen on ultrasound examinations. In addition, primary hyperparathyroidism was evaluated by screening individuals for hypercalcemia.

The work was conducted in two stages. The first was a Pilot Study, the primary purpose of which was to evaluate the feasibility of the methods proposed, and to develop the specific operational procedures and data collection instruments needed for a full study. The second stage was to implement the remaining fieldwork to complete such a study. This approach allowed the accumulation of information and experience prior to initiation of the more costly full-scale study.

The Pilot Study was completed in December 1994, with a report issued January 24, 1995. Reviews of the Pilot Study by the National Research Council's Board of Radiation Effects Research of the Commission on Life Sciences and the federal Advisory Committee for the HTDS concluded that a full-scale epidemiologic study should be undertaken. The fieldwork for the Full Study was completed in December 1997. This document summarizes the Final Report of the Hanford Thyroid Disease Study.

II. Fieldwork

A. Cohort Definition and Participant Selection

To achieve the primary objective of the study, it was important to identify a cohort that would provide the greatest likelihood of detecting an association between Hanford radiation exposure and thyroid disease, if such a relationship exists. This was accomplished by defining a cohort that would include adequate numbers of people with the highest possible radiation doses to the thyroid from Hanford, as well as people with very low radiation doses to the thyroid from Hanford.

Extensive efforts were made to investigate different sources of information that could be used to construct a cohort of people who might have been exposed. Ideally, such a list would include everyone in a relatively large population living in the region around the Hanford site during the time period that atmospheric releases of radioactive iodine occurred, and would contain enough identifying information on each person to allow them to be located for the study (several decades after exposure). Only birth records provided a viable unbiased source for identifying a cohort.

For the purposes of participant selection only, residence at time of birth was considered a surrogate for the anticipated radiation dose to the thyroid from Hanford, since doses could only be estimated from data collected during the study. To select study participants for the Pilot Study, a birth roster was constructed based on all births to mothers resident in the Washington State counties of Benton, Franklin, Walla Walla, Okanogan, Ferry, and Stevens. Following the Pilot Study, and based on the dose estimates for Pilot Study participants, Adams County was added for the Full Study selections, in order to maximize the numbers of participants with high doses. Adequate numbers of participants with no or very low dose were obtained in the Pilot Study selections from Stevens, Ferry, and Okanogan counties, and no further selections were made from these areas.

Preliminary estimates from the HEDR project suggested that the highest thyroid doses were likely to be in people exposed as infants or children during the first years of Hanford operations. This is because infants and children receive higher thyroid doses per unit exposure, due primarily to the small size of their thyroid glands. Existing literature also suggests that the risks radiation-induced thyroid disease (and possibly hyperparathyroidism) are greatest among those exposed at youngest ages. For these reasons, the Pilot Study included people born from 1942-46, since the large majority of atmospheric releases of radioactive iodine from the Hanford facility occurred in 1944-46. For the Full Study, additional selections from the years 1940 and 1941 in Benton, Franklin, and Adams counties were included to maximize the number of potentially high dose participants. Thus, the cohort contained people with exposure beginning as early as the prenatal period, and as late as age three. An additional benefit of choosing a young group was that mothers and close relatives of participants born during 1940-46 were more likely be alive and available for interview, compared to those of persons born earlier.

Selection of potential participants from the Birth Roster was stratified by geographical area, year of birth, and sex. The purpose of stratification by geographical area and birth year was to assure that adequate numbers of high dose and low dose participants were identified, so that as wide a range of doses as possible was obtained. Stratification by sex also reduced the possibility of confounding by sex that could reduce the efficiency of the study. Geographical areas were defined to distinguish predominantly rural areas from those that are predominantly urban, because residents of predominantly rural areas may have been more likely to consume fresh raw milk than their more urban counterparts. A total of 5199 individuals were selected to form the cohort.

B. Tracing and Locating Study Participants

Because members of the study cohort were identified solely on the basis of birth records from the mid-1940s, extensive effort was required to locate them as adults nearly fifty years later. Thus, the primary objective of the tracing component of the study was to identify a current address and telephone number for all living cohort members, so they could be recruited to participate in the study. A second objective was to obtain confirmation of death, as well as date and cause of death, for all those who were deceased.

Several approaches were used to trace potential participants. Initially, relatively easy to use and readily accessible sources were used. Subsequently, more resource-intensive and costly resources were employed to find the more difficult to locate individuals. A final attempt to locate the most difficult to find potential participants was made by using established professional locating services and military locating services.

Of the 5199 cohort members, 4350 living individuals were located and 527 individuals were confirmed deceased. Thus, nearly 94% of the cohort was located, with their identities confirmed. Only 322 potential participants (6.2%) remained "unable to locate" at the end of the study. Notably, the ability to locate well over 90% of all potential participants did not vary substantially by sex, or geographic region of birth, or year of birth. Almost 84% of all potential participants were located as living, and their identities (whether they agreed to participate or not) were confirmed directly by contact with the potential participants themselves or with close relatives who could verify their identities and current addresses.

Five hundred twenty-seven (10.1%) of the cohort members were confirmed to be deceased by a close relative and/or other reliable source (such as death certificate). The proportion confirmed deceased was higher among males (12.7%) than females (7.5%). Sixteen potential participants (0.3%) were located as living, but died during the study prior to completing a clinic. Death certificates were obtained for 93% of the total 543 deceased.

At least one living cohort member was located in every state in the U.S. except for Rhode Island. Fifty-four percent of those located resided in Washington State, 9.4% in California, 9.1% in Oregon and 2.7% in Idaho. The only other state where more than 2% of the living cohort members resided was Texas (2.2%). Thirty-six participants (0.8% of those located) lived in countries outside of the U.S., including Canada, Dubai, Ecuador, Germany, Mexico, Saudi Arabia, South Africa, South Korea, England, Guam, Australia, Japan, France, Saipan, Hungary, Columbia, and Taiwan.

C. Recruiting Study Participants

The objectives of the recruiting effort were to contact living cohort members, obtain their agreement to participate in the study, and to identify an appropriate respondent to complete the Computer Assisted Telephone Interview (CATI). Once a potential participant was located through the tracing procedure, initial contact was made by mail. In some instances a preliminary letter or phone call was necessary to confirm the potential participant's identity. Each living potential participant located received an initial contact letter, fact sheet, and a description of what participation in the study would entail.

A recruiter called each located potential participant five to seven days after the first contact letter was mailed. A minimum of 10-15 evening attempts were made at various weeknight and weekend time periods, and a minimum of three daytime (weekend and weekday) calls were attempted. If the potential participant could not be contacted by phone after 20-25 attempts, a second letter was sent explaining that the study had been unable to reach them at the phone number on file, and asking them to call the toll-free HTDS number. After 40-45 attempts resulting in no contact with either the potential participant or a household member, the potential participant was considered “unable to contact” and no further attempts were made.

If a potential participant refused, the recruiter asked him/her to complete a Refusal/Demographic Questionnaire. Twelve demographic questions relating to race, ethnic origin, income, religion, and education level were asked in order to obtain a general profile of those who refused to participate, or who withdrew after initially agreeing to participate. The recruiter also completed a Refusal Assessment after the call to record the nature and strength of the refusal from the recruiter's perspective.

A total of 4239 potential participants (97.4% of all living, located cohort members) were contacted by telephone and invited to participate in the study. An additional 93 (2.1% of all living, located cohort members) were located to an address, and were sent one or more letters, but could not be contacted by telephone. Of those contacted by telephone, 3564 (84.1%, or 81.9% of all located, living cohort members) agreed to participate in the study. Of those located alive, 634 (14.6%) refused to participate.

Willingness to participate did not differ substantially by sex, year of birth, or geographic region of birth. "Not interested" and/or "no time" were by far the most commonly given reasons given refusals, accounting for 64.8% of all refusals. The second most commonly cited reasons were "illness" and "impairment" (7.6%). An additional 41 potential participants were determined to be unable to fully participate during the recruiting process and were consequently not included in the study regardless of willingness to participate.

D. Computer Assisted Telephone Interview

The primary objective of the Computer-Assisted Telephone Interview (CATI) was to collect information that would be used as input for calculating an estimated radiation dose to the thyroid gland for each study participant. A CATI was conducted by an interviewer who read the interview text and questions from a computer screen, and recorded the responses as they were given.

The CATI was designed to collect information from the early years of the participants' lives, including time *in utero* if necessary, from 1944 to 1957. The interview was "location-driven" so that the information collected was specific to locations and periods of time directly relevant to the atmospheric releases of ^{131}I from Hanford. The following topic areas were included in the CATI interview: 1) general demographic characteristics of the participant and his or her family; 2) a residential history of the participant from birth through 1957, and for the mother while pregnant with or breastfeeding the participant; 3) sources of the milk consumed by the participant from birth through 1957, and by the participant's mother while pregnant and breastfeeding the participant; 4) milk consumption patterns of the participant from birth through 1957, and of the mother during pregnancy and breastfeeding; and 5) other patterns of food consumption, including green and leafy vegetables, fruit, and free range chicken eggs by the participant from birth through 1957, and by the mother while pregnant and breastfeeding. In addition, medical history information was obtained for both the mother and the participant, including the following: 1) thyroid diseases and selected other medical conditions diagnosed and treated in the participant; and 2) history of medical radiation exposures, either diagnostic or therapeutic, for the participant, and for the mother during pregnancy and breastfeeding.

To help the CATI respondents accurately report detailed information about their child (or sibling) from very long ago, several elements of the cognitive approach to interviewing were incorporated into the design of the CATI. The key element to this approach is to mentally take the respondent back to the time period in question, and have them remember as much about that time as possible. As more memories of the time in question are recalled by the respondent, the likelihood of remembering answers to specific questions increases.

Memory materials were developed to help the respondent prepare for answering the interview questions. Background information was provided to encourage memory about specific topics. The memory

materials were organized into a booklet that was sent with a Residence History Questionnaire to respondents in advance of the interview. In addition, the text of the interview was refined to include references to specific parts of the memory materials at key points during the interview.

Of the 2712 participants who identified a CATI respondent, interviews were completed for 2266 (83.6%). Of the 3447 eligible study participants who completed the clinic, 2133 (61.9%) had a CATI interview. In 29 instances, CATI interviewers determined the quality of the data provided by respondents was too poor to be considered reliable. Expanded interviews were performed at the clinic for these 29 participants.

E. Scheduling

The primary objective of the scheduling activity was to provide each participant with at least three options for clinic attendance, with the least possible inconvenience to the participant. A schedule of clinic dates and locations was developed based on the current residences of participants. Clinics were held in Seattle, Pasco, Spokane, Walla Walla, Yakima, Wenatchee, Colville, Omak, Portland Oregon and Vancouver Washington. Most participants from outside Washington State attended clinics in Seattle.

Multiple attempts were made to contact all participants, and every participant was offered several options for clinic dates. Each scheduled participant was sent a letter that included: 1) the date and time of clinic appointment; 2) location of clinic and directions; 3) travel arrangements summary and/or tickets (if applicable); and 4) Interview Preparation Worksheet. If a participant canceled a clinic appointment, an attempt was made to reschedule the participant. A participant who canceled a clinic appointment would be rescheduled an unlimited number of times. If a participant decided not to participate in the study during the scheduling process, the scheduler assessed the reason for the withdrawal and addressed the participant's concerns in an attempt to retain participation. If the participant persisted in the withdrawal, she or he was asked to complete a Refusal Questionnaire.

Approximately 90% of those who initially agreed to participate completed a clinic. The number of participants who withdrew after initially agreeing to participate was 298 (7.7%).

F. Clinical Evaluation

The objective of the clinical component of the study was to provide a thorough clinical examination of each study participant to determine the presence of thyroid disease, or primary hyperparathyroidism. Each participant was administered an In-Person Interview prior to the clinic examinations; this is described in more detail below. Following the interview, each participant underwent a full complement of examinations to determine the presence or absence of any thyroid disease or primary hyperparathyroidism. The examinations included thyroid ultrasound, independent thyroid palpation by two study physicians, and blood tests for thyroid and parathyroid function, and anti-thyroid immune response. Additional studies were requested if indicated by the presence of palpable thyroid nodules.

The physical examination was conducted separately by two study physicians. The results of their examinations were reviewed, and if there was any disagreement, the two examiners conferred and re-examined the participant together to reach a consensus. The findings of each physician were recorded separately, as were the findings of any consensus examination, prior to review of the ultrasound scan. If abnormalities were found on the ultrasound which were not found on physical exam, the two physicians performed a final consensus examination. The physical examination and ultrasound findings were then discussed with the participant.

Participants found to have discrete, palpable, solitary thyroid nodules or dominant nodules within a multinodular gland upon examination were asked to undergo fine-needle aspiration (FNA) biopsy of the nodule. Participants who wanted to delay the procedure could either return to the HTDS clinic site on another clinic date, or have the FNA performed by a local physician in their community. Thyroid nuclear scans were recommended for participants whose examination and laboratory results were suspicious for the presence of autonomously functioning thyroid nodules, Graves Disease, or toxic thyroid nodules.

A total of 3447 eligible participants were examined in the HTDS clinics. Of the 3447 participants, 3439 (99.8%) had blood drawn for thyroid function studies, and 3446 had thyroid ultrasound. Of the 272 participants for whom FNA was recommended, 259 (95.2%) underwent the procedure, while 28 of the 29 (96.6%) participants recommended to have a nuclear scan complied.

G. In-Person Interview

The purpose of the In-Person Interview was to obtain information directly from the study participant about his/her past exposures to occupational and/or medical irradiation, history of thyroid disease, and general demographic information. In addition, for those participants who could not identify a respondent for the dosimetry interview, an expanded version of the In-Person Interview provided details regarding residence history and limited information on the type of milk consumed, for use in estimating their thyroid radiation doses from Hanford's atmospheric releases of ^{131}I. The In-Person Interview was conducted before the participant began the medical components of the clinic (ultrasound, blood draw, and physical examination). This was done to ensure that the participant's responses could not be influenced by knowledge of examination results. All interviews were conducted in person by trained, experienced interviewers.

The In-Person Interview included questions about the participant from age 15 to the present, in the following topic areas: 1) general demographic characteristics; 2) residential history, including dates and locations of residences, 3) occupational history, focusing on occupations and industries with potential of exposure to any form of ionizing radiation; 4) military history as obtained in both the residential and occupational sections, especially regarding possible exposures to nuclear weapons tests (e.g., in Nevada or the Marshall Islands); 5) medical history, including dates and places for all thyroid-related diseases and symptoms; 6) history of medical and dental X-ray exposures; 7) history of nuclear medicine procedures; 8) history of radiation therapy; 9) selected lifestyle factors, such as patterns of tobacco use; and 10) familiarity/bias questions to assess knowledge of the Hanford releases and any strongly-held beliefs about their possible health effects.

All 3447 eligible participants attending a HTDS clinic completed an In-Person Interview. Six interviews were judged to have insufficient residence history information to calculate a dose estimate. One participant was unable to complete the interview because of developmental disabilities, however the participant's father (who was unable due to illness to participate in a CATI dosimetry interview) was mailed a modified version of the expanded In-Person Interview questionnaire and provided the information in this manner. Overall, 61% of participants completed the Standard In-Person Interview, while 39% completed the Expanded version.

H. Clinic Medical Review and Final Diagnosis Determination

The objectives of the Clinic Medical Review and Final Diagnosis Determination processes were to: 1) evaluate each participant's clinical thyroid examination results from the HTDS clinic visit; 2) communicate clinic results to participants in a timely manner and, when indicated, to the participant's health care provider; and 3) assign the final diagnoses for each case, according to the format developed using all information available prior to and including the HTDS clinic.

Following each clinic, results of the laboratory tests performed on blood specimens, of radiologists' reviews of ultrasound examinations, and of the study pathologist's evaluation of any FNA specimens were received in the HTDS office within 5-6 days. Physicians reviewed each participant's clinic results, and a letter informing the participant of the results was sent. A Final Diagnosis Determination Form was completed for all remaining participants. All participants received their clinic results within 3-4 weeks following their clinic appointment. Letters were also sent to each participant's health care provider, if the participant indicated this was to be done.

If follow-up tests were recommended to a participant, that participant's clinic and follow-up results were reviewed at another Clinic Medical Review once the results were received in the HTDS office. A second results letter was mailed to the participant and their health care provider, describing the results of the follow-up tests. The Final Diagnosis Determination Form was then completed.

All 3447 eligible participants who attended a study clinic received a Clinic Medical Review. Eighty percent of participants had a Final Diagnosis Determination Form completed at the time of their Clinic Medical Review. The remaining 20% had either historical medical records or post-clinic recommendations for further diagnostic procedures, and had a Final Diagnosis Determination Form completed following compilation and review of the records from those providers.

A total of 259 participants had FNA procedures performed at the clinic or on the recommendation of the HTDS physicians. Of these, 47 were recommended at Clinic Medical Review to have further biopsy or surgical procedures to rule out a diagnosis of thyroid neoplasm. In addition, 29 participants with thyroid nodules or suppressed TSH were recommended to undergo thyroid nuclear scan. Twenty participants had an abnormal calcium level and were recommended to have additional blood drawn for parathyroid hormone (PTH) studies to confirm or rule out a diagnosis of hyperparathyroidism. Thirty participants were requested to have additional blood drawn due to abnormal or borderline thyroid function.

I. Historical and Post-Clinic Medical Records Review

The primary objectives of the medical record component were to: 1) document thyroid problems reported by study participants and CATI respondents; 2) obtain any cytological or histological specimens from previous biopsies or surgeries for review by the study's pathologist; and 3) obtain the results (including histological specimens) of any further diagnostic or surgical procedures recommended by the HTDS as a result of a finding at the HTDS clinic. A secondary objective of the medical record component was to obtain cause of death information on all deceased cohort members, in order to assign cause of death codes and perform a mortality analysis.

During the CATI interview, respondents were asked to provide the names (and addresses, if known) of any physician who saw the participant for diagnosis or treatment of thyroid disease. At the time of the In-Person Interview, the participant was asked to provide the names and addresses of physicians or institutions where they had been diagnosed or treated for thyroid or parathyroid disease, and to sign a consent form for the release of information from each of these providers.

For each deceased cohort member, the death certificate or informant information was used to complete a Cause of Death Form. In addition, the primary cause of death was coded using the ICD9-CM system. For those whose date of death preceded the use of the ICD9-CM system, the primary cause of death was also back-coded using the system in use at the time of death.

Reports of historical medical records were obtained for 694 participants, with a total of 1259 consent forms completed to obtain medical records from different providers. While the majority of reports were made during the In-Person Interview, CATI Interviews yielded 30 of these reports.

Of the 1259 Medical Record Consents obtained, a total of 795 (63.1%) separate medical records were received by the HTDS. No records were received for 464 requests (36.9%). In 102 (8.1%) cases, records could not be requested because the physician was deceased or retired, or a current address could not be identified. For 128 (10.2%) requests, records were unavailable due to the destruction of records, the inability of the provider to identify the patient, or an inability to locate the records. In 232 (18.4%) cases, records were not received after several contacts, without explanation as to why they were not available.

Of the 694 participants identifying historical medical records to be requested, pathology or cytology slides were requested for 52 (7.5%). In a few cases, more than one set of slides was requested, for a total of 58 separate requests. A total of 42 sets of historical pathology or cytology slides were received for 42 (80.8% of slides requested) participants.

One potential concern is that diagnoses of disease outcomes might be missed when requested medical records or slides could not be obtained: none or only part of the requested records or slides were received for 199 (29%) and 160 (23%), respectively, of the 694 participants for whom such requests were made. However, even if a medical record or slide could not be obtained, the likelihood of a missed diagnosis was generally low because in most such situations the HTDS evaluation provided a definitive assessment of whether the diagnosis for which the medical record was sought was confirmed or not confirmed.

Medical records documenting further diagnostic studies recommended as a result of the HTDS clinic findings were requested for 35 participants, with a total of 72 separate requests. All but one of these records were obtained, with at least one record obtained for each of the 35 participants. Thirty-three of these participants also had histology or cytology slides requested, for a total of 35 separate requests. All 35 of these specimens were obtained.

Death certificates were received for 504 of the 543 deceased cohort members. Cause of death was coded for 543 deceased cohort members.

J. Dose Estimation

The primary analyses of dose-response relationships were based on individual estimates of radiation dose to the thyroid from Hanford's atmospheric releases of ^{131}I, specifically organ doses to the thyroid that were estimated from data collected during the CATI and/or the Expanded In-Person Interview. The CIDER program developed by the HEDR Project was used to calculate estimated doses. In particular the CIDER output for an individual consisted of 100 realizations of the estimated cumulative total organ dose to the thyroid from ^{131}I.

Each of the 100 realizations of dose was calculated for a fixed set of conditions regarding the source term, environmental transport, and uptake of ^{131}I, and these conditions for a given realization were the same for every participant. The 100 realizations were obtained by randomly varying the conditions, i.e., the uncertain parameters in the HEDR models for source term, transport, etc., in order to characterize the uncertainty in the resulting dose estimates. Thus it is useful to view each realization as consisting of a set of doses, one for each in-area participant. For many purposes it was useful to have a single number or

"point estimate" to represent each participant's dose. For each living evaluable in-area participant, the median of the 100 realizations of dose, $d_i = \text{median}(D_{i,1}, \ldots, D_{i,100})$ for participant i, was used as a summary measure of that participant's dose. .

Of the 3447 eligible participants who attended a study clinic, 3440 were considered evaluable for the study, i.e., had sufficient information for dose estimation and could be adequately examined for thyroid disease. The CIDER program calculated estimates of doses accumulated by people while living within a 75,000 square mile geographical domain around Hanford. Dose estimates could therefore be calculated by the CIDER program for 3191 of the evaluable participants who lived within that domain at least some time from the start of Hanford operations in 1944 through the end of 1957; these 3191 are designated "in area" participants. The remaining 249 "out-of-area" participants did not, according to their CATI or Expanded In-Person Interview data, live within the domain during that time period. Although the CIDER could not calculate dose estimates for the out-of-area participants, they were included in the study.

K. Data Management

The primary objective of the Data Management Plan was to specify the procedures to develop and maintain the study databases, and the procedures that would be used to ensure data quality. Principal components of the plan included duplicate entry for all data forms, range checks encoded in the data entry programs, and consistency check programs run on the data after entry. A second objective of the Data Management Plan was to maintain the confidentiality of the data. This included data in computerized form through the use of passwords and control of limited access to directories and data files, and to paper records, which were stored in locked files in locked offices or in a file room which had limited access via keycard.

In order to ensure high data quality, all data entered from paper forms were subject to double-entry verification. Additional computer programs were written to check and crosscheck all of the data, both within a data form and across data forms. For example, the diagnoses coded on the Final Diagnosis Determination Forms were compared to all the other data collected (i.e., examination forms, ultrasound forms, CATI data, In-Person Interview data, and the tracking system) to ensure that all appropriate diagnoses were included. All inconsistencies were investigated by review of the participant's records, including audiotapes of CATI interviews when necessary. Once any changes were made to a database, check programs were run to ensure all changes had been made correctly.

L. Data Quality Control

In addition to the data management plans and procedures outlined above, additional steps were taken after data collection to ensure a high degree of data quality. These efforts included 1) more extensive between-table consistency checks of the In-Person Interview data and the CATI data, 2) hand calculation of the participant's diet portion of the CIDER input data files ("scenario files") for 10% of those with a CATI, 3) comparison of the mother's diet portion of the scenario file for all those with a CATI based on a separate computer program written by a programmer other than the one who created the scenario file program, 4) comparison of dose estimates produced by a CDC programmer versus those produced by HTDS, and 5) review by a second programmer of complex analysis programs that included code other than standard SAS procedures.

III. Special Considerations

A. Coordination with the Advisory Committee

In June of 1990, an Advisory Committee was appointed by the Secretary of the Department of Health and Human Services to advise and consult with the CDC regarding the design and conduct of the study. The committee was established pursuant to the *Federal Advisory Committee Act, (5 U.S.C. Appendix 12)*. The role of the committee was to review the development of the study protocol and conduct of the Pilot Study, assist in determining the feasibility and design of a full-scale epidemiologic study, and advise CDC on the analysis of the study results.

Initially, meetings of the committee were to be held on a quarterly basis in Atlanta, Georgia. In recognition of the interest in the Pacific Northwest in such proceedings, however, the committee asked that at least one meeting per year be held in Washington State. Following completion of the Pilot Study, meeting frequency was reduced to approximately once per year, with the majority of these held in Seattle, Washington.

Meetings of the Advisory Committee were open to the public. All materials presented to the committee became public record, with copies available for members of the public at the meetings. Meetings held in Washington State were nearly always accompanied by an evening Public Meeting to allow members of the public to attend and to ask questions or make comments regarding the study.

Each meeting of the Advisory Committee began with an update on the progress of the study since the previous meeting. These presentations included the status of preparations for the study field work, or later, the numbers of study participants completing each phase of the study. Updates on the separate work concerning Native American populations were also included. Requests for further information from the committee were documented, and information provided by study staff and investigators, as necessary.

B. Public Involvement

An important aspect of the HTDS was the provision of prompt, accurate, and complete information to the public. In this context it was crucial that contacts be established with members of the populations most interested in the work. Interested parties included representatives of the States of Washington, Oregon, and Idaho; the Native American Tribes and Nations in the study areas, and local area residents.

The public information activities of the study were designed to accomplish the following goals: 1) to assure that residents of the region understood the issues that led to the initiation of the study, the purpose and objectives of the study, its basic epidemiologic design, and the time schedule within which it was to be conducted; 2) to provide opportunities for the public to express concerns and comments regarding the design and conduct of the study, and to answer public questions regarding all aspects of the project; 3) to create public interest and support for the study, particularly in ways that such support might enhance participation by persons selected to be study participants; and 4) to assure broad dissemination and proper interpretation of final study results.

Throughout the study, and particularly in the early phases, the study investigators participated in public meetings held during the bi-monthly meetings of the HEDR Technical Steering Panel (TSP), and contributed to the planning activities of the Communications subcommittee of the TSP. The HTDS also supplied the TSP with a fact sheet that was included with TSP fact sheet mailings. This written material was updated periodically as the study progressed.

Several separate approaches were also taken to provide information to the public regarding the HTDS. Initially, the study protocol was made available for public review and comment prior to its submission to the CDC and the Advisory Committee. In conjunction with this activity, a series of public meetings were held throughout the Northwest to discuss the protocol with the public and to answer specific questions.

In addition to the study fact sheet mentioned above, several study brochures were developed and a newsletter describing the progress and status of the study was initiated. The brochures included the following: 1) HTDS Fact Sheet; 2) Questions and Answers about the Study; 3) Questions and Answers about Radiation and Thyroid Disease; and 4) Review of Thyroid Disease and Approach to Diagnosis. A master mailing list, which included the lists previously maintained by the FHCRC, the CDC, and the HEDR Project, was assembled to mail the newsletter and brochures to interested individuals. By the later stages of the HTDS, the mailing list contained nearly 9700 names. Early in the study, the newsletter was published on a quarterly basis. Following the Pilot Study, however, yearly updates were used to inform interested parties of the study's progress. A total of 15 issues were published. A special issue summarizing the findings in the Draft Final Report was distributed in January 1999.

Finally, study investigators and staff have been available to answer questions on a regular basis. A phone line was designated in the Seattle study office for public inquiries, and a toll-free telephone number was established at the Fred Hutchinson Cancer Research Center for the Hanford Thyroid Disease Study (1-800-638-HTDS). People selected as study participants, and members of the general public, were encouraged to use the toll-free number to contact the study office if they had questions or scheduling conflicts. As access to the World Wide Web became more common, a web site for the study was established at the FHCRC. All study brochures and newsletters have been available at that site since January 1997, and are updated as necessary.

C. Native American Component

Nine Native American tribes and nations have reservations and ceded lands in the region around Hanford: Colville, Couer d'Alene, Kalispell, Kootenai, Nez Perce, Spokane, Umatilla, Warm Springs, and Yakama. Members of these tribes and nations were exposed to ^{131}I from Hanford, and the original congressional mandate that led to the HTDS called specifically for the inclusion of "Indian tribes and tribal organizations." The approach taken in the HTDS regarding the Native American populations was determined by two important characteristics of those populations. First, the lifestyles of many Native Americans were quite different in many respects from those of the non-Native population. In particular, many Native Americans followed traditional cultural practices, especially regarding diet and sources of foods, which might influence the doses they received from Hanford's ^{131}I, but which were not explicitly modeled in the CIDER program. Moreover many Native Americans maintained a seasonal migratory pattern of residence. Second, because the tribes and nations have sovereign rights recognized by the United States, conduct of a research project such as HTDS would require the approval and active cooperation of each tribal government. Thus, the objective of the HTDS with respect to the Native American populations was to assess the feasibility of conducting a study to determine whether thyroid disease has increased among Native Americans exposed to atmospheric releases of ^{131}I from Hanford.

Sample size and power calculations were carried out to determine whether it would be feasible to conduct a retrospective cohort study using individual dose estimates, similar to that being conducted for the HTDS Full Study. These calculations were based on data provided by eight of the nine tribes regarding tribal-specific lifestyle and dietary practices. These data are likely to more accurately account for lifestyle patterns and practices specific to each tribe than using assumptions derived from the non-Native American population, and therefore the representative dose estimates are likely to more accurately approximate the doses members of each tribe would have likely received from Hanford. Similarly, the demographic data provided by each tribe are likely to more accurately reflect the size and demographic makeup of each tribe around the time of the Hanford releases.

Even under very liberal assumptions regarding the number of tribal members who might be available to participate in a study, and the thyroid radiation doses Native Americans received from Hanford's atmospheric releases of ^{131}I, a study nearly twice the size of the HTDS Full Study (6426 living evaluable participants) would have only 50% power to detect an effect of the magnitude targeted by the Full Study, i.e., a 5% increase in total thyroid neoplasia per Gray. Even under a more extreme assumption that the baseline probabilities for thyroid neoplasia are only half of those assumed in the HTDS Full Study, a study of 6426 living evaluable participants would only have 71% power to detect the same magnitude of effect. Thus, it was recommended that a study of the design of the HTDS full study would not be feasible in the Native American population encompassed by the nine tribes in the Hanford region.

IV. Statistical Methods

In the primary dose-response analyses, the exposure for each living evaluable in-area participant was represented by the estimated radiation dose to the thyroid from Hanford's atmospheric releases of ^{131}I, as calculated using the CIDER program created by the HEDR Project. The primary dose-response analyses for disease outcomes and ultrasound-detected abnormalities (UDAs) of the thyroid were based on regression models in which the probability of having the outcome of interest varies as a linear function of estimated thyroid dose, specifically the median dose as mentioned above. The model for this primary analysis permitted the background probability of the outcome to depend on sex, but assumed a common regression coefficient (slope) for the dose-response. The regression coefficient can be interpreted as the change in the probability of the disease outcome, per unit change in dose. Since the purpose of the study was to determine whether thyroid disease has been increased, significance testing focused on the null hypothesis that the probability of having the outcome of interest does not vary with dose (i.e., that the regression parameter has value zero) and the one-sided alternative hypothesis that the probability increases with increasing dose (i.e., that the regression parameter is greater than zero). Alternative sex-stratified dose-response models were also considered, specifically linear-quadratic and logistic models.

Identification and analysis of confounding and effect modifying factors was accomplished through the analysis of generalizations of the logistic exposure-response models. For disease outcomes, these generalizations allowed the background probabilities of the outcome of interest (i.e., the intercept parameters) and/or the regression parameters to vary as functions of a number of factors that might potentially confound the relationship between thyroid radiation dose and the outcome of interest. The influence of uncertainty of the dose estimates on the dose-response relationships was examined by 1) fitting the linear dose-response model using each of the 100 realizations of dose separately, and 2) using a Bayesian approach to calculate deattenuated estimates of the regression slope parameter in the sex-stratified logistic model.

It was not assumed that the out-of-area participants were unexposed to ^{131}I from Hanford. Indeed, results of the HTDS Pilot Study suggested that many out-of-area participants lived in locations near the HEDR domain at various times during 1945-1957. Alternative methods of assigning a dose estimate for out-of-area participants were developed, and these dose estimates were used to assess the sensitivity of dose-response results to assumptions about the doses for out-of-area participants.

The distribution of doses was quite skewed, with large numbers of comparatively low doses and small numbers of quite high doses. Therefore analyses were performed to assess whether the dose-response results might be inordinately influenced by the high dose participants. In particular, two empirical checks were made to assess whether the estimated regression coefficient adequately represents the dose-response relationship over the lower dose range.

Information released by the U.S. National Cancer Institute (NCI) shortly before and during October 1997, indicated that people living in the contiguous 48 states during the 1950s and 1960s were exposed to various levels of ^{131}I released from the Nevada Test Site (NTS). The material released by NCI included estimates of dose for representative individuals in all counties in the 48 states, as well as more detailed data regarding estimated dose by individual test detonation, county, and age. Limited preliminary comparisons for HTDS participants suggested that in many cases the reported NTS dose estimates were comparable to or even greater than the estimated Hanford doses. Therefore it was judged necessary to evaluate exposure to ^{131}I from the NTS as a potential confounding factor. For each participant in the HTDS, the "estimated NTS dose" was defined specifically as the thyroid dose from ^{131}I entering the atmosphere from tests conducted at NTS between 1951 and 1957, inclusive, as estimated from data made publicly available by NCI. A categorical variable representing each living evaluable participant's relative level of exposure to ^{131}I from the NTS was calculated for use in the analyses of potential confounding.

V. Summary of Dose-Response Results

The primary evaluation of dose-response relationships focused on twelve categories of thyroid disease, hyperparathyroidism, and ultrasound-detected abnormalities of the thyroid. For each of these 14 outcome categories a primary case definition was specified based on the most definitive diagnostic criteria available. Diagnostic information obtained from the HTDS evaluation and diagnostic information which was well documented in medical records and met criteria for HTDS diagnoses was considered to be the most definitive and of the highest quality. The primary analysis for each outcome was therefore restricted to cases defined according to these two sources. The principal dose-response analysis used this primary definition of outcome, individual radiation dose estimates (the median of CIDER's 100 realizations for each individual) based on individual residence history, dietary consumption data from the CATI when available, and HEDR default values when such data were not available. The results from these analyses using the primary outcome definition constitute the principal findings of the HTDS.

Additional criteria were also defined for each outcome category, to identify cases using less definitive diagnostic criteria, e.g., information obtained from prior medical records that did not meet HTDS criteria, or reports of a diagnosis by a participant or CATI respondent which could not be confirmed by the HTDS evaluation or medical records. Although the principal findings of the HTDS are based on the primary outcome definition, dose-response analyses were also conducted for each of these alternative definitions with less definitive diagnostic criteria. In addition, dose-response analyses were conducted for six outcome categories based on the results of laboratory assays, and for thyroid mass estimated from the ultrasound scan. Dose-response analyses for all disease and thyroid UDA outcomes were repeated using two alternative sets of individual dose estimates, and two alternative representations of exposure that did not use the CIDER program to estimate individual radiation doses. Efforts were also made to evaluate the influence of uncertainties in individual dose estimates on the fitted dose-response relationships for the primary case definition in each outcome category.

In overall summary of the dose-response results, there was no evidence of a statistically significant association between estimated thyroid radiation dose from Hanford and the cumulative incidence of any of the 14 primary outcomes. There was also no evidence of any statistically significant dose-response relationship for any of the alternative definitions of outcome. The findings were essentially unchanged for analyses based on either of the two alternative sets of individual dose estimates. The results remained the same after taking into account several factors that might confound the relationship between radiation dose and the outcome of interest. There was no evidence of any statistically significant dose-response for any outcome that might be different from the linear model used in the primary analyses (e.g., a linear-quadratic or logistic relationship). Incorporation of uncertainty in the dose estimates did not significantly change the primary results for any of the outcomes or change the overall conclusions of the study. Summarized below are the main findings for each of the primary outcomes investigated.

Thyroid Cancer

Twenty (0.6%) of the 3440 living evaluable participants were diagnosed with thyroid cancer; 13 women (0.7%) and 7 men (0.4%). In all but one case, the diagnosis was based on histologic evidence from the HTDS examination (12) or prior histologic evidence (7). Using the primary definition (19 total cases; 14 in-area) and maximum likelihood analysis of the sex-stratified linear probability model, the cumulative incidence of thyroid cancer did not increase significantly with estimated dose ($p = 0.25$), with an estimated slope of 0.002 per Gy, and Bonferroni-adjusted 95% confidence interval (CI) ranging from less than –0.001 to 0.017 per Gy. Analyses that considered less definitive criteria to identify cases and alternative dose estimates or representations of exposure revealed no statistically significant dose-response relationships. Incorporation of uncertainty in the dose estimates did not significantly change the primary results.

Benign Thyroid Nodule

Two hundred and forty-nine (7.2%) of the 3440 living evaluable participants had a diagnosis of benign thyroid nodule based on histologic or cytologic evidence arising from the HTDS examination or from a prior diagnosis which met HTDS diagnostic criteria; 170 (9.7%) women and 79 (4.7%) men. An

additional 38 participants (1.1%) had HTDS or prior diagnoses classified as clinical (i.e., palpable nodule with no cytology or histology available), and another 10 (0.3%) had diagnoses based solely on a report by the participant or his/her CATI respondent. Using the primary definition (249 total cases; 235 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the cumulative incidence of benign thyroid nodule did not increase significantly with estimated dose ($p = 0.68$), with an estimated slope of –0.008 per Gy, and Bonferroni-adjusted 95% CI ranging from less than –0.022 to 0.041 per Gy. Analyses that considered less definitive criteria to identify cases, as well as other disease outcomes related to benign nodules (e.g., benign nodules and nodules suspicious for follicular neoplasm, benign nodule excluding non-neoplastic disease, solitary nodule detected without ultrasound, benign nodule excluding colloid-only nodules, and benign colloid nodules), and analyses which considered alternative dose estimates or representations of exposure, revealed no statistically significant dose-response relationships. Accounting for potential confounding or effect modification, and incorporation of uncertainty in the dose estimates, did not significantly change the primary results.

Total Thyroid Neoplasia

This outcome was defined to include participants with thyroid cancer based on HTDS or prior histology, or benign thyroid nodule with a histologic type of follicular adenoma, based on HTDS or prior histology. A total of 33 (1.0%) of the 3440 living evaluable participants were included in this category; 20 (1.1%) women and 13 (0.8%) men. Using the primary definition (33 total cases; 28 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the cumulative incidence of total thyroid neoplasia did not increase significantly with estimated dose ($p = 0.42$), with an estimated slope of 0.001 per Gy, and Bonferroni-adjusted 95% CI ranging from less than –0.003 to 0.022 per Gy. Analyses using alternative dose estimates or representations of exposure revealed no statistically significant dose-response relationships. Incorporation of uncertainty in the dose estimates did not significantly change the primary results.

Any Thyroid Nodule

This outcome was defined by the diagnosis of one or more of the following: benign thyroid nodule, thyroid cancer, or nodule suspicious for follicular neoplasm. A total of 281 (8.2%) of the 3440 living evaluable participants had this outcome based on histologic or cytologic evidence arising from the HTDS examination or from a prior diagnosis; 193 (11.0%) women and 88 (5.2%) men. Another 39 (1.1%) were based on HTDS or prior clinical diagnoses (i.e., palpable nodule with no available cytology or histology), and 10 living evaluable participants had diagnoses of any thyroid nodule based solely on reports from the participant or his/her CATI respondent.

Using the primary definition (281 total cases; 261 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the cumulative incidence of any thyroid nodule did not increase significantly with estimated dose ($p = 0.65$), with an estimated slope of –0.007 per Gy, and Bonferroni-adjusted 95% CI ranging from less than –0.023 to 0.043 per Gy. Analyses that considered less definitive criteria to identify cases and alternative dose estimates or representations of exposure revealed no statistically significant dose-response relationships. Accounting for potential confounding or effect modification, and incorporation of uncertainty in the dose estimates, did not significantly change the primary results.

Hypothyroidism

Two hundred and sixty-seven (7.8%) of the 3440 living evaluable participants had a diagnosis of hypothyroidism based on the HTDS evaluation or on medical records with supporting documentation; 204 (11.7%) women and 63 (3.7%) men. An additional 105 (3.1%) living evaluable participants had a diagnosis of hypothyroidism based on medical records but without supporting documentation, and 30 (0.9%) were inferred from past or current thyroxine therapy. This latter group consisted of participants who were taking thyroid hormone replacement, but in whom no medical records were available to confirm the original diagnosis of hypothyroidism. There were also 193 (5.6%) cases based solely on reports of hypothyroidism from the participant or his/her CATI respondent.

Using the primary definition (267 total cases; 246 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the cumulative incidence of hypothyroidism did not increase significantly with estimated dose ($p = 0.61$), with an estimated slope of –0.006 per Gy, and Bonferroni-adjusted 95% CI ranging from less than –0.016 to 0.047 per Gy. Analyses which considered less definitive criteria to identify cases, as well as permanent hypothyroidism, and analyses which considered alternative dose estimates or representations of exposure, revealed no statistically significant dose-response relationships, although the estimated regression coefficients from logistic regression analyses using less definitive criteria to identify cases were somewhat larger. Accounting for potential confounding or effect modification, and incorporation of uncertainty in the dose estimates, did not significantly change the primary results.

Autoimmune (Hashimoto's) Thyroiditis

A total of 625 (18.2%) of the 3440 living evaluable participants had a diagnosis of autoimmune thyroiditis based on the HTDS evaluation or medical records with supporting documentation; 403 (23.1%) women and 222 (13.1%) men. Another three cases were based on medical records without supporting documentation, and one case was based solely on a report by the participant or his/her CATI respondent.

Using the primary definition (625 total cases; 582 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the cumulative incidence of autoimmune thyroiditis did not increase significantly with estimated dose ($p = 0.82$), with an estimated slope of –0.026 per Gy, and Bonferroni-adjusted 95% CI ranging from less than –0.057 to 0.044 per Gy. Analyses which considered less definitive criteria to identify cases, additional outcomes related to the assay for antithyroid immune response, and autoimmune thyroiditis in combination with non-iatrogenic, permanent hypothyroidism, as well as analyses which considered alternative dose estimates or representations of exposure, revealed no statistically significant dose-response relationships. Accounting for potential confounding or effect modification, and incorporation of uncertainty in the dose estimates, did not significantly change the primary results.

Graves Disease

A total of thirty-four (1.0%) of the 3440 living evaluable participants had a diagnosis of Graves Disease based on the HTDS evaluation or on medical records with supporting documentation; 28 (1.6%) women and 6 (0.4%) men. Three (0.1%) living evaluable participants had a diagnosis of Graves Disease based on medical records without supporting documentation, and an additional thirteen (0.4%) were based solely on a report from the participant or his/her CATI respondent.

Using the primary definition (34 total cases; 32 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the cumulative incidence of Graves Disease did not increase significantly with estimated dose ($p = 0.56$), with an estimated slope of –0.001 per Gy, and Bonferroni-adjusted 95% CI ranging from less than –0.002 to 0.024 per Gy. Analyses that considered less definitive criteria to identify cases and alternative dose estimates or representations of exposure revealed no statistically significant dose-response relationships. Accounting for potential confounding or effect modification, and incorporation of uncertainty in the dose estimates, did not significantly change the primary results.

Autoimmune Thyroid Disease

Autoimmune thyroid disease was defined by a diagnosis of autoimmune (Hashimoto's) thyroiditis or Graves disease based on the HTDS evaluation or medical records with supporting documentation. A total of 659 (19.2%) of the 3440 living evaluable participants were included in this category; 431 (24.7%) women and 228 (13.5%) men. These included 625 with autoimmune (Hashimoto's) thyroiditis and 34 others with diagnoses of Graves disease. An additional 4 (0.1%) living evaluable participants had a diagnosis of autoimmune thyroid disease based on medical records without supporting documentation

(three with autoimmune thyroiditis, one with Graves disease). Eleven others (0.3%) were based solely on a report by the participant or his/her CATI respondent (one with autoimmune thyroiditis, 10 with Graves disease).

Using the primary definition (659 total cases; 614 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the cumulative incidence of autoimmune thyroid disease did not increase significantly with estimated dose ($p = 0.80$), with an estimated slope of –0.024, and Bonferroni-adjusted 95% CI ranging from less than –0.058 to 0.048 per Gy. Analyses that considered less definitive criteria to identify cases and alternative dose estimates or representations of exposure revealed no statistically significant dose-response relationships. Accounting for potential confounding or effect modification, and incorporation of uncertainty in the dose estimates, did not significantly change the primary results.

Hyperthyroidism

A total of 161 (4.7%) of the 3440 living evaluable participants were diagnosed with hyperthyroidism based on the HTDS evaluation or medical records with supporting documentation; 134 (7.7%) women and 27 (1.6%) men. An additional 14 (0.4%) living evaluable participants had a diagnosis of hyperthyroidism based on medical records without supporting documentation, and 21 (0.6%) were based solely on a report from the participant or his/her CATI respondent. It is important to note that these 196 cases included a substantial number of iatrogenic cases caused by excess thyroid hormone replacement. Since endogenous hyperthyroidism (hyperthyroidism not caused by thyroid hormone over-replacement) was of particular importance, analyses that focused on cases of non-iatrogenic hyperthyroidism were emphasized in this study.

Using the primary definition (161 total cases; 155 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the cumulative incidence of hyperthyroidism did not increase significantly with estimated dose ($p = 0.22$), with an estimated slope of 0.011 per Gy, and Bonferroni-adjusted 95% CI ranging from less than –0.008 to 0.052 per Gy. Analyses that considered less definitive criteria to identify cases, as well as non-iatrogenic hyperthyroidism, and analyses which considered alternative dose estimates or representations of exposure, revealed no statistically significant dose-response relationships. Accounting for potential confounding or effect modification, and incorporation of uncertainty in the dose estimates, did not significantly change the primary results.

Multinodular Thyroid Gland

A total of 95 (2.8%) of the 3440 living evaluable participants had a diagnosis of multinodular thyroid gland based on the HTDS evaluation; 73 (4.2 %) women and 22 (1.3 %) men. An additional nineteen (0.6%) living evaluable participants had a diagnosis of multinodular thyroid gland based on medical records, and one diagnosis was based solely on a report from the participant or his/her CATI respondent.

Using the primary definition (95 total cases; 85 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the cumulative incidence of multinodular thyroid gland did not increase significantly with estimated dose ($p = 0.88$), with an estimated slope of –0.006 per Gy. The lower limit of the Bonferroni-adjusted 95% confidence interval was not estimated, but the upper limit was 0.014 per Gy. Analyses that considered less definitive criteria to identify cases and alternative dose estimates or representations of exposure revealed no statistically significant dose-response relationships. Accounting for potential confounding or effect modification, and incorporation of uncertainty in the dose estimates, did not significantly change the primary results.

Simple Goiter

The diagnosis of simple goiter was uncommon, with only 14 (0.4%) of the 3440 living evaluable participants having this diagnosis based on HTDS evaluation; 9 (0.5%) women and 5 (0.3%) men. Another

28 (0.8%) had diagnoses based on medical records, and for an additional 28 (0.8%) the diagnosis was based solely on a report by the participant or his/her CATI respondent.

Using the primary definition (14 total cases; all in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the cumulative incidence of simple goiter did not increase significantly with estimated dose ($p = 0.74$), with an estimated slope of –0.001 per Gy. The lower limit of the Bonferroni-adjusted 95% confidence interval was not estimated, but the upper limit was 0.012 per Gy. Analyses that considered less definitive criteria to identify cases and alternative dose estimates or representations of exposure revealed no statistically significant dose-response relationships. Incorporation of uncertainty in the dose estimates did not significantly change the primary results.

Other Thyroid Disease

Four living evaluable participants, all in the in-area group, had diagnoses of other thyroid disease based on their HTDS examinations or medical records with supporting documentation. These included two cases of subacute thyroiditis, one case of familial thyroglobulin binding deficiency, and one case of secondary hypothyroidism. The first alternative definition added only two cases of subacute thyroiditis with diagnoses based on medical records without supporting documentation. For both the primary and first alternative definition of other thyroid disease, there were too few cases for meaningful estimation of the radiation dose-response.

The second alternative definition added 20 participants, primarily with participant or CATI respondent reports of past thyroid disease of unknown type. This brought the total number of cases to 26, of whom four were out-of-area participants. Based on maximum likelihood analysis of the sex-stratified linear probability model using this case definition, the estimated slope was slightly greater than zero (0.002 per Gy), with Bonferroni-adjusted 95% confidence interval ranging from less than –0.002 to 0.024 per Gy, providing no evidence that the cumulative incidence increased significantly with increasing dose (one-tailed $p = 0.39$). Because the number of cases in this category was small, and the diagnoses were heterogeneous and mostly unknown, further analyses of this outcome were not performed.

Hyperparathyroidism

A total of 12 (0.3%) living evaluable participants had a diagnosis of hyperparathyroidism based on the HTDS evaluation or on medical records with supporting documentation; 10 (0.6 %) women and 2 (0.1%) men. Another two diagnoses were based on a report from the participant or his/her CATI respondent. One additional living evaluable participant who did not meet the study's criteria for hyperparathyroidism nevertheless had an elevated calcium level in the presence of a high normal PTH level, when the PTH should have been suppressed, highly suggestive of hyperparathyroidism. This participant was included as a case in an additional analysis.

Using the primary definition (12 total cases; 11 in-area), the cumulative incidence of hyperparathyroidism did not increase significantly with estimated dose ($p = 0.61$), with an estimated slope of –0.0001 per Gy. The lower limit of the Bonferroni-adjusted 95% confidence interval was not estimated, but the upper limit was 0.013 per Gy. Analyses that considered less definitive criteria to identify cases and alternative dose estimates or representations of exposure revealed no statistically significant dose-response relationships. Incorporation of uncertainty in the dose estimates did not significantly change the primary results.

Ultrasound-Detected Abnormalities of the Thyroid (Thyroid UDAs)

The thyroid gland was visible in the ultrasound examinations of 3429 of the 3440 living evaluable participants. For 11 participants the thyroid was not visible, 10 because of thyroid surgery and one because the sonographer couldn't adequately visualize the thyroid. Among the 3429 whose thyroids were visible, 1596 (46.5%) had one or more thyroid UDAs; 964 (55.5 %) women and 632 (37.4 %) men. Ultrasound findings were categorized as palpable thyroid UDAs (224 or 6.5%), nonpalpable focal thyroid UDAs (1309

or 38.2%), and diffuse thyroid UDAs (458 or 13.4%). All three types of UDA were more frequent among women than men. Ultrasound-detected thyroid abnormalities were based only on the HTDS evaluation, not on any prior ultrasound scans.

Based on maximum likelihood analysis of the sex-stratified linear probability model, the prevalence of any UDA (1596 total cases; 1481 in-area) did not increase significantly with estimated dose ($p = 0.21$), with an estimated slope of 0.031 per Gy, and Bonferroni-adjusted 95% CI ranging from –0.059 to 0.116 per Gy. Similarly, the prevalence of palpable UDA (224 total cases; 204 in-area) did not increase significantly with estimated dose ($p = 0.95$), with an estimated slope of –0.018 per Gy. The Bonferroni-adjusted lower 95% confidence limit was not estimated due to the magnitude of the negative slope estimate, however the upper confidence limit was 0.015 per Gy. The prevalence of nonpalpable focal UDA (1309 total cases; 1217 in-area) also did not increase significantly with estimated dose ($p = 0.23$), with an estimated slope of 0.027 per Gy and Bonferroni-adjusted 95% CI ranging from –0.061 to 0.115 per Gy. Analyses of all three types of ultrasound abnormalities in relation to alternative dose estimates or representations of exposure revealed no statistically significant dose-response relationships. Accounting for potential confounding or effect modification, and incorporation of uncertainty in the dose estimates, did not significantly change the primary results.

Additional analyses were performed to assess whether ultrasound abnormalities might be precursors to radiation-induced clinical disease. These analyses evaluated whether increasing dose was associated with increasing prevalence of large thyroid UDAs, increasing number of thyroid UDAs, or the presence of diffuse thyroid UDAs. To assess whether the dose-response results might be affected by the size of focal thyroid UDAs, three additional outcomes were analyzed. These included the presence of a focal UDA with maximum dimension at least 5 mm, the presence of a focal UDA with maximum dimension at least 10 mm, and the presence of a focal UDA with average dimension at least 15 mm. These additional analyses applied only to palpable and nonpalpable focal thyroid UDAs, since diffuse UDAs were not defined by any size criterion. In none of these additional analyses was there any evidence that the risk of having a focal UDA of a particular size increased with increasing dose ($p=0.64$, 0.88 and 0.53 for the presence of focal UDA with maximum dimension of 5 mm, maximum dimension of 10 mm and average dimension of 15 mm, respectively).

To assess whether the number of thyroid UDAs detected in individual participants might increase in relation to estimated thyroid radiation dose, the numbers of focal thyroid UDAs with maximum dimension ≥ 5 mm, maximum dimension ≥ 10 mm, and average dimension ≥ 15 mm were counted for each living evaluable participant with an HTDS ultrasound examination. No statistically significant dose-response was found between estimated thyroid radiation dose and the average number of focal thyroid UDAs ($p = 0.80$, 0.48 and 0.43 for the number of thyroid UDAs meeting the three size criteria, respectively).

The prevalence of diffuse thyroid UDA (458 total cases; 428 in-area) did not increase significantly with estimated dose ($p = 0.14$), with an estimated slope of 0.029 per Gy, and Bonferroni-adjusted 95% CI ranging from –0.029 to 0.100 per Gy. Analyses that considered alternative dose estimates or representations of exposure revealed no statistically significant dose-response relationships. Accounting for potential confounding or effect modification, and incorporation of uncertainty in the dose estimates, did not significantly change the primary results.

Laboratory Tests and Thyroid Mass

Of the 3191 living evaluable in-area participants, 3183 (99.7%) provided a blood specimen at the HTDS clinic. Several laboratory assays were conducted to evaluate thyroid function, anti-thyroid antibody response, and serum calcium level. In addition to the dose-response analyses conducted of specific thyroid disease outcomes, which incorporated information from these tests in the determination of the diagnosis, dose-response analyses were also conducted to investigate whether there were associations between the laboratory values from these tests and estimated thyroid radiation dose from Hanford (i.e., regardless of thyroid disease diagnosis).

Thyroid stimulating hormone (TSH) levels were measured for all participants who provided a blood specimen. Of the 3183 living evaluable in-area participants who provided blood samples, 222 were receiving exogenous thyroid hormone at the time of their HTDS clinic and were excluded from the analyses of TSH. Among the remaining 2961 living evaluable in-area participants, three different TSH assays were used during the study. There was no statistically significant trend of average TSH level in relation to estimated radiation dose for any of the three assays considered either separately or in a combined analysis. Free thyroxine index (FTI) was analyzed, also excluding the 222 participants who were receiving exogenous thyroid hormone at the time of their HTDS clinic. There was no significant trend of FTI in relation to estimated radiation dose (two tailed $p = 0.23$). Three different tests for antithyroid antibody response (anti-TPO, AMA, and anti-TG) were used over the course of the study. There was no significant trend of any assay result in relation to estimated radiation dose ($p = 0.66$ for anti-TPO, 0.52 for AMA, and 0.20 for anti-TG).

Serum calcium levels were measured in an effort to identify participants with hypercalcemia that might be secondary to hyperparathyroidism. Of the 3183 living evaluable in-area participants who provided blood samples, 227 with diagnoses of hypothyroidism or hyperthyroidism based on the HTDS examination were excluded from the primary analysis of serum calcium levels. Two additional participants did not have serum calcium data due to insufficient volumes of collected blood. There was a statistically significant trend of decreasing serum calcium level in relation to increasing radiation dose ($p = 0.0074$), with an estimated slope of -0.09 per Gy, and Bonferroni-adjusted 95% CI ranging from -0.16 to -0.01 per Gy. Although there is no readily apparent explanation for this result, this finding deserves further comment. First, the outcome for which calcium was being measured, hyperparathyroidism, was not found to be associated with radiation dose. Second, the dose effect occurred within the normal range of calcium values. For both women and men, the estimated background means were about $9.2 \pm .01$, consistent with the normal range of the test (8.4 to 10.2). Only 0.9% of the cohort had low calcium levels less than 8.4 (hypocalcemia). There was no statistically significant relationship between hypocalcemia and radiation dose. Third, even at a dose of 3000 mGy to the thyroid, which is larger than the maximum estimated dose for any study participant, the mean serum calcium levels predicted by the regression model were well within the normal range. Therefore, despite the *statistically* significant decrease in calcium levels with increasing dose, the resulting effect or clinical impact does not appear to be *clinically* significant.

Estimates of thyroid mass were available for 3400 living evaluable participants for whom both lobes of the thyroid were visible on ultrasound; 3153 were in-area participants. There was no statistically significant trend of thyroid mass in relation to estimated radiation dose ($p = 0.98$).

VI. Summary Comments and Conclusions

The HTDS was conducted to determine whether exposure to atmospheric releases of radioactive iodine, in particular ^{131}I, from the Hanford Nuclear Site between 1944 and 1957 resulted in increased thyroid disease among those exposed. The study evaluated twelve categories of thyroid disease, hyperparathyroidism, ultrasound-detected abnormalities of the thyroid, the results of several laboratory tests for thyroid function, anti-thyroid antibody and serum calcium level, and thyroid mass. The primary analysis (based on HTDS diagnostic criteria of the highest level of certainty) utilized an estimate of thyroid radiation dose for each individual based on information about their residence history and dietary consumption patterns during the times of the Hanford releases. Additional analyses were conducted using several alternative methods for estimating dose, both quantitative and qualitative, including methods that were independent of the HEDR models. The primary analyses were based on a linear dose-response model, adjusting for the effects of differences in response by sex, although alternative models for the shape of the dose-response were also investigated. The potential confounding or dose-response modifying effects of a number of lifestyle factors and indicators of other radiation exposure were evaluated. All primary dose-response analyses were repeated to include adjustments for uncertainty in the individual radiation dose estimates.

This study found no statistically significant association between dose to the thyroid from Hanford radiation and 1) cumulative incidence of any of the disease outcomes; 2) prevalence of ultrasound-detected thyroid abnormalities; or 3) thyroid laboratory tests or thyroid mass. There was also no statistically significant dose-response for hyperparathyroidism, although increasing thyroid dose was significantly associated with a decrease in average serum calcium level. Although the explanation for this result is not clearly apparent, the finding does not appear to be of clinical significance. These results remained the same when alternative methods of assessing radiation dose were used, and after accounting for uncertainty in dose estimation. Based on data available regarding the tracing and enrollment of study participants, there is no evidence that the absence of a dose-response relationship is due to bias in selection of the cohort, loss to follow-up, or enrollment and participation.

Although no statistically significant dose-response was found for any of the disease outcomes in this study, many study participants had thyroid disease. A considerable effort was made to assess the world literature on the prevalence of the major thyroid and parathyroid disease outcomes evaluated in the HTDS. Studies selected for review were those conducted in other locations and most comparable to the HTDS for the outcomes of thyroid nodules, thyroid cancer, hypothyroidism, autoimmune thyroiditis, hyperparathyroidism, and thyroid UDAs. This was done in order to compare the disease experience of the HTDS cohort to what might reasonably be expected based on the experience in other populations not exposed to Hanford radiation. As discussed in the Section X.E of the Report, comparisons of this type are imperfect and must be interpreted with great caution. Differences in prevalence estimates between the HTDS cohort and other populations may well reflect differences in any of a number of factors other than exposure to radiation from Hanford. Nevertheless, from review of these studies, it appears that estimates of cumulative incidence derived from the HTDS are well within the range and are consistent with published estimates. There is no indication that the levels of thyroid or parathyroid disease occurrence in the HTDS cohort are systematically different, or higher, than what has been reported around the world in a variety of different circumstances.

Given the differences between the radiation exposure circumstances at Hanford and those of other populations in which radiation-related risks of thyroid disease have been studied, the findings of this study are not inconsistent with the current published literature regarding the effect of exposure to ^{131}I and the risk of thyroid and parathyroid disease. This is particularly so given the relatively small magnitude of the estimated thyroid radiation doses in HTDS study participants (mean = 174 mGy) and the relatively protracted nature of the exposure over time. There is little evidence in the literature to suggest that people exposed to ^{131}I at the levels found in this study over a period of months or years would experience higher rates of thyroid or parathyroid disease as a result of their exposure.

Nevertheless, a lingering question for many may be whether the uncertain nature of the dose estimation used in the primary analyses is so great that it renders the quantitative dose-response results

inconclusive. The study has attempted to address this possibility in three ways. First, alternative qualitative methods of assigning exposure were used. Results from these analyses were consistent with those from the quantitative dose-response analyses. Second, two different approaches were employed to evaluate the impact of dose uncertainty on the primary risk estimates. Neither resulted in findings that were significantly different from those ignoring such uncertainty. Third, the impact of dose uncertainty on study power was assessed using simulation methods. These analyses revealed that the reduction in statistical power due to uncertainty in dose estimation was modest, and that even after accounting for such uncertainty the study had adequate statistical power to detect effects as small or smaller than those in the existing published literature. Although any epidemiologic study is limited to some extent by uncertainty in the assessment of exposure, the impact of such uncertainty on the power of the study and the estimation of risk is seldom addressed to the extent attempted here. Further, the fact that epidemiologic investigations are inherently "uncertain" does not imply complete randomness or unpredictability, nor does it mean that reasonable conclusions cannot be drawn from such studies.

In conclusion, the results of the HTDS provide no evidence of a statistically significant association between increasing thyroid radiation dose from Hanford and the cumulative incidence of any of the primary outcomes studied. These findings do not definitively rule out the possibility that Hanford radiation exposures are associated with an increase in one or more of the outcomes under investigation. However, it does mean that if such associations exist, they were likely too small to detect using the best epidemiologic methods available.

I. INTRODUCTION

The Hanford Thyroid Disease Study (HTDS) was conducted by a team of investigators at the Fred Hutchinson Cancer Research Center (FHCRC) under contract to the Centers for Disease Control and Prevention (CDC) Radiation Studies Branch. The Study Management Team (SMT) which had primary responsibility for the design and conduct of the investigation, consisted of Scott Davis, Ph.D. (epidemiology), Kenneth Kopecky, Ph.D. (biostatistics), and Thomas Hamilton, M.D., Ph.D. (endocrinology). Bruce Amundson, M.D. (family medicine) was a member of the SMT through August 1998. Ms. Peggy Adams Myers served as Project Manager through the release of the Draft Final Report. Ms. Beth King assumed responsibility for project management thereafter. In addition to the FHCRC team, the study employed Dr. Robert Griep as an expert consultant on thyroid disease. Dr. Bruce Kulander served as the pathologist who reviewed all pathological specimens. Four radiologists at Seattle Nuclear Medicine/Ultrasound Associates interpreted the thyroid ultrasound scans. Administrative, statistical, and technical staff reported directly to the Project Manager and the SMT. The clinical component of the HTDS was directed by Dr. Hamilton, with the assistance of the HTDS study physicians in conducting thyroid examinations. Study operations were based at the FHCRC in Seattle, with a field office in the Tri-Cities for the Subject Tracing component.

The CDC was kept informed on a monthly basis of progress in the design and conduct of the study, and provided technical support as needed by the FHCRC. Mr. Michael Sage and Mr. Michael Donnelly served as the Project Officers. Dr. Paul Garbe was the primary scientific liaison. In addition, an Advisory Committee was appointed for this study by the Secretary of the Department of Health and Human Services to provide advice and consultation to the CDC and the SMT.

The technical approach to this research project was divided into three phases. The first phase involved the development of the study protocol and preparation for the Pilot Study. These preparations included the appointment and convening of the Advisory Committee and approval of the protocol by the federal Office of Management and Budget (OMB) and the Institutional Review Board (IRB) of the FHCRC. This phase began upon award of the contract in September 1989, and was concluded in late 1992. The other two phases of the study, the Pilot Study and the Full Study, are discussed further below.

The study was conducted as a follow-up prevalence study. That is, potential participants were selected on the basis of presumed past level of exposure to radioactive iodine from Hanford, based on place and year of birth. Participants were located and evaluated for the presence or history of thyroid disease. Information was also collected regarding each participant's residence and dietary history in order to estimate his or her thyroid radiation dose from Hanford. The primary analyses focused on living participants who received medical examinations to detect thyroid disease, and for whom individual thyroid radiation doses could be estimated using the dosimetry system developed by the Hanford Environmental Dose Reconstruction (HEDR) Project and the information collected by the HTDS. Although the effects of primary interest are defined by three categories of thyroid disease (hypothyroidism, benign thyroid nodules, and thyroid cancer), information regarding all forms of thyroid disease were recorded as part of the study and are included in the overall analysis. In addition, hyperparathyroidism was evaluated by screening individuals for hypercalcemia. Since the aim of the study was to investigate whether risks of the thyroid diseases were increased by exposure to Hanford's ^{131}I, the analysis examined whether the cumulative incidence of these diseases increased with increasing dose to the thyroid.

The methods of the study can be summarized as follows. Potential study participants were selected from birth records to form a cohort for follow-upPeople likely to have lived in a seven-county geographic area surrounding the Hanford Site were selected to ensure as much as possible that the cohort contained people with a wide range of radiation doses to the thyroid (e.g., from the highest doses to very low doses). Attempts were made to trace and locate each individual in the cohort. Once located, each person was invited to a medical clinic for a thorough diagnostic evaluation for thyroid disease. At the clinic, each study participant: 1) underwent a personal interview regarding details of his/her residential,

medical, and personal histories; 2) provided a blood sample for thyroid function tests, antibody markers for autoimmune thyroiditis, and serum calcium determination; 3) received a thyroid ultrasound examination; and 4) received a physical examination of the thyroid by two physicians independently of one another. For those found to have palpable thyroid nodules or nonpalpable ultrasound detected thyroid nodules ≥ 1.5 cm (average of three dimensions), permission was sought to conduct a fine needle aspiration to provide more complete diagnostic information. To verify reports of thyroid diseases that occurred in the past, medical records and pathology specimens were sought and reviewed in a uniform manner.

Prior to the participant's clinic visit, an attempt was made to interview the mother, or other close relative knowledgeable about aspects of the participant's childhood that influenced the radiation dose he or she received from Hanford. The information collected in this interview was used to estimate radiation dose to the thyroid using algorithms developed by Battelle Pacific Northwest Laboratory as part of the Hanford Environmental Dose Reconstruction (HEDR) Project. Detailed descriptions of each component of the study fieldwork are found in section V of this report.

Following the development and approval of the study protocol, the research was conducted in two subsequent phases. The first was a Pilot Study. The primary purpose of this phase was to evaluate the feasibility of the methods proposed and to develop the specific operational procedures and data collection instruments needed for a Full Study. Once the results of the Pilot Study indicated that it was feasible to conduct a successful full-scale epidemiologic study, the second stage was implemented to complete the remaining fieldwork for the Full Study. This approach allowed the accumulation of information and experience prior to initiation of the more costly full-scale study. This also allowed for the possibility that the design and procedures for the Full Study could be modified if necessary to account for the realities of the field environment.

Eleven Pilot Study objectives were specified in the original HTDS protocol (1). These objectives dealt with both logistical and statistical issues. Logistical issues to be evaluated included: 1) the efficacy and success rates of the fieldwork procedures, including the use of birth certificates to identify potential study participants; 2) the ability to trace and locate persons identified; 3) the ability to collect information for use in estimating thyroid radiation dose; 4) the success in bringing participants to clinics for thyroid examinations and 5) the costs of these activities. Statistical issues to be evaluated included: 1) estimating the distributions of radiation dose to the thyroid among groups of individuals defined by place of birth; 2) evaluating the suitability of the areas from which participants were selected to ensure a cohort of individuals with a full range of doses; and 3) calculation of the statistical power that could be reasonably achieved in a Full Study and the sample size required to do so. A detailed report of the results of the Pilot Study was submitted to the CDC on January 24, 1995. A summary of the findings of that report is included here as Appendix 1 (Executive Summary of the Pilot Study Report).

It should be emphasized that in testing the feasibility of the study design, it was important to evaluate procedures and instruments for participants who were likely to have received high doses from Hanford radiation releases, as well as for those who were not likely to have received such doses. There was concern that the degree to which individuals could be identified, traced, located, and recruited into the study might be influenced to a large extent by their physical proximity to the Hanford Site and their perception of any direct threat to their own health from Hanford. Thus, a very important aspect of the Pilot Study was to evaluate the success of including people who lived in varying proximity to the Hanford Site.

It is also important to emphasize that the Pilot Study was not designed to assess health outcomes in relation to radiation dose. Instead, the Pilot Study was designed to: 1) test the feasibility of the proposed field logistics; 2) estimate the radiation doses likely to have been received among study participants and, therefore, to determine the distribution of doses according to factors such as geographical area (e.g., urban vs. rural), age, and sex; and 3) derive the information necessary to adequately plan a Full Study that would be capable of determining whether radiation releases from Hanford resulted in an increased risk of thyroid disease or hyperparathyroidism. The number of participants included in the Pilot Study was too small and the individual radiation dose estimates available from the HEDR Project were too preliminary to enable

any formal evaluation of adverse health effects in the pilot phase of the HTDS. Thus, no estimates of thyroid disease or hyperparathyroidism risk associated with exposure to radioactive iodine were reported at the conclusion of the Pilot Study. However, all data obtained from individuals who participated in the Pilot Study were included in the Full Study.

The Pilot Study was completed in late 1994. To maintain study operations in anticipation of conducting the Full Study, it was necessary to define a "transition phase" between the Pilot Study and Full Study. In the fall of 1993 the Federal Advisory Committee and the CDC gave approval to select an additional sample of 1000 potential participants to serve as a Transition Sample. Based on the information available at that time from the Pilot Study, it was decided that the Full Study would likely be implemented. Thus, the Transition Sample was selected to enable field operations to continue while the Pilot Study was completed and its results evaluated. This approach also shortened the time to complete the Full Study.

A report of the results of the Pilot Study was prepared and submitted to the CDC and the Advisory Committee in January 1995. Each of the objectives outlined for the Pilot Study in the study protocol was evaluated. The report's major conclusions were that:

1. The thyroid dose distributions obtained in the Pilot Study, which were the basis for the sample size and power calculations, were reasonably representative of what the overall dose distribution would be at the completion of the Full Study.

2. To achieve sufficient statistical power to detect an increase of 5% in thyroid neoplasia per Gray, it would be necessary to enroll a minimum of approximately 3200 living evaluable participants.

3. The basic design and data collection methods would remain the same.

4. Estimation of doses study participants would be conducted by HTDS staff by remote access to the HEDR computer programs at the CDC in Atlanta.

5. All births from the following years to mothers living in the indicated counties should be added to the cohort:
   a) 1942-1944: Remaining Richland, Pasco/Kennewick and Benton County
   b) 1940-1941: All of Benton and Franklin Counties
   c) 1940-1944: All of Adams County

The primary criterion for continuing with the Full Study was the ability to identify and recruit adequate numbers of people with a sufficient range of radiation doses. Specifically, the aim was to design the Full Study to have statistical power of at least 90% to detect a linear dose-response for the probability of having thyroid neoplasia (malignant or benign) with a slope of 0.0001 per rad (10% per Gy). If the results of the Pilot Study had indicated that it would not be possible to obtain at least 80% power to detect an effect of this magnitude, then consideration would have been given to terminating the study. However, the results of the Pilot Study (2) revealed that, not only did the procedures and plans work well for all aspects of the study, a conservative projection of statistical power of 80% to detect an increased risk of thyroid neoplasia of 5.0% per Gray was possible with some revisions to geographic areas and years of birth sampled. Section V.A. of this report discusses in detail the sampling utilized in the Full Study to achieve this level of power.

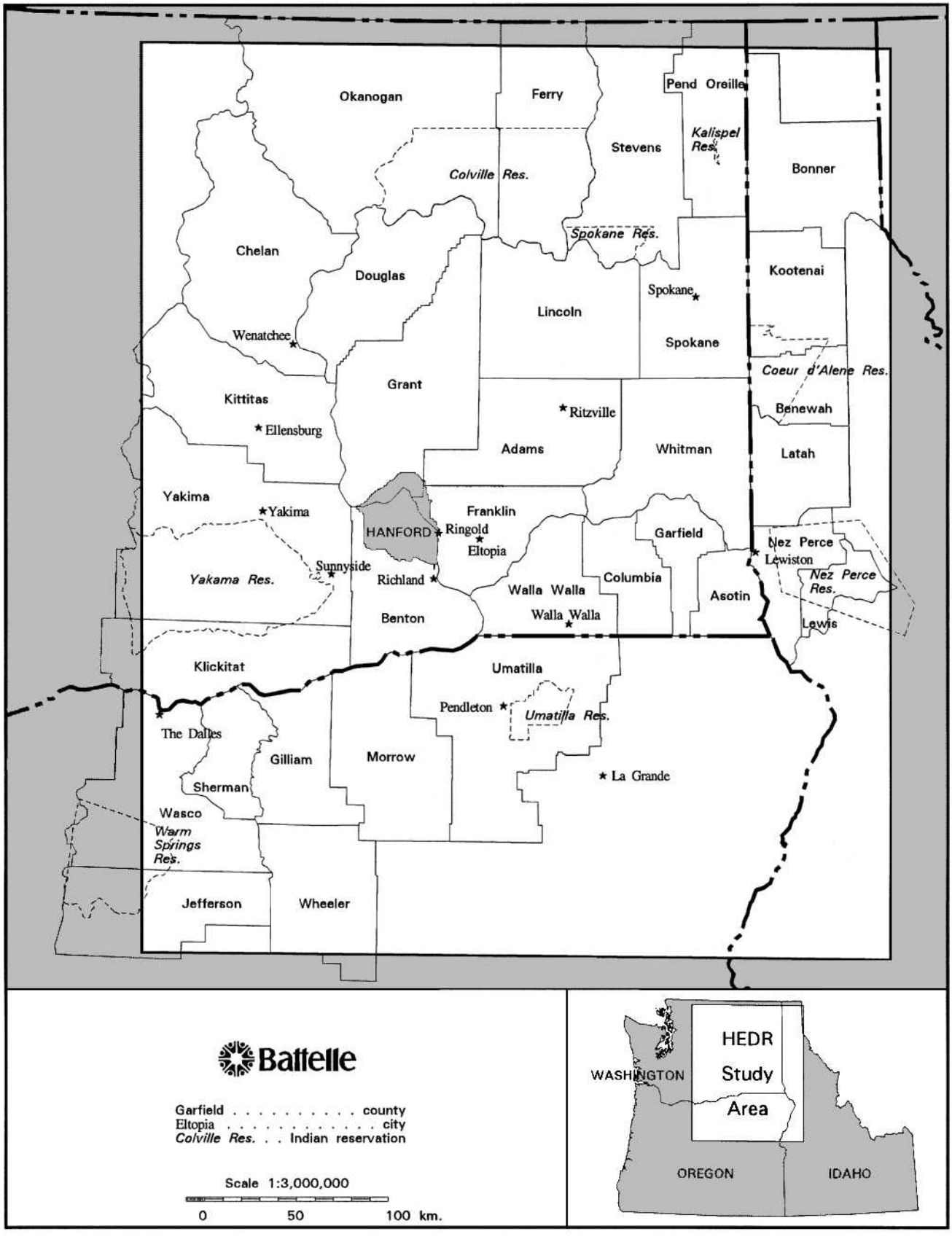
Thus, in February 1995, the Advisory Committee recommended to the CDC that the Full Study be done. The data collection phase of the Full Study was completed in late 1997, and was followed by a period of data analysis and the reporting of results. The purpose of this report is to document in detail the conduct of each phase of the HTDS, and the results of the analyses that were done.

II. BACKGROUND

A. Historical Perspective

The Hanford Nuclear Site occupies an area of approximately 560 square miles in southeastern Washington adjacent to the towns of Pasco, Kennewick, and Richland (Figure II.A-1). The facility was established in 1943 as part of the Manhattan Project to produce plutonium-239 for the development of the first nuclear weapons. The initial Hanford production reactor (B) became operational in September of 1944, and was followed by a second (D) in December. Two chemical separations plants, which constituted the second phase of plutonium production, began to operate in 1944 and 1945. By 1957 there were six additional production reactors and two fuel reprocessing plants on the Site.

Figure II.A-1 HEDR Study Area


Okanogan
Ferry
Pend Oreille
Kalispel Res.
Stevens
Colville Res.
Bonner
Spokane Res.
Chelan
Douglas
Kootenai
Spokane
Lincoln
Spokane
Wenatchee
Coeur d'Alene Res.
Grant
Kittitas
Ritzville
Benewah
Ellensburg
Adams
Whitman
Latah
Yakima
Franklin
Yakima
HANFORD
Ringold
Eltopia
Garfield
Nez Perce
Lewiston
Sunnyside
Yakama Res.
Richland
Columbia
Nez Perce Res.
Walla Walla
Asotin
Benton
Walla Walla
Lewis
Klickitat
Umatilla
Pendleton
Umatilla Res.
The Dalles
Gilliam
Morrow
La Grande
Sherman
Wasco
Warm Springs Res.
Jefferson
Wheeler
Battelle
Garfield . . . . . county
Eltopia . . . . . city
Colville Res. . . . Indian reservation
Scale 1:3,000,000
0
50
100 km.
HEDR
Study
Area
WASHINGTON
OREGON
IDAHO

The creation of the Hanford Nuclear Site caused the regional population in the Columbia Basin to expand rapidly. Although the original construction force was large (approximately 50,000 persons), most had left the area by the end of 1945. However, major Hanford expansions in the late 1940s resulted in substantial population growth in the Tri-Cities area of Richland, Pasco, and Kennewick. Between 1945 and 1960, the population increased from approximately 40,000 to over 100,000 (3). In addition, Army anti-aircraft units numbering about 5000 troops per year served at Hanford beginning in 1950. Army personnel and construction workers and their families lived in a trailer-barracks enclave about five miles north of Richland. The counties surrounding the Hanford site, traditionally ranching and agricultural areas, continued to be populated by small, family farms. With the establishment of the Columbia Basin Irrigation Project and subsequent agricultural development, large numbers of new families moved into the region in the late 1950s.

In February 1986, largely as the result of repeated public requests from the Hanford Education Action League (Spokane, Washington) and the Environmental Policy Institute (Washington, D.C.), as well as requests from the Centers for Disease Control and Prevention and the State of Washington, the Department of Energy made available over 19,000 pages of documents (many of which were previously classified) describing radiation releases and environmental monitoring during the early years of plutonium production at Hanford. Based on information found in these documents and a subsequent Freedom of Information Act (FOIA) request, an additional 20,000 pages were made public in April 1987. Approximately 25,000 more pages have been released since late 1987.

Data contained in this material indicate that during the initial years of plutonium production at Hanford substantial quantities of radionuclides were released into the atmosphere, particularly during the first few years of production. In attempting to produce plutonium rapidly in 1944 and 1945, irradiated uranium was allowed to decay approximately 45 days before being treated for reprocessing. As a result, the subsequent chemical treatment of the irradiated uranium produced large atmospheric releases of gaseous radionuclides. A primary component of these releases was a radioactive isotope of iodine, iodine-131 (^{131}I).

Although it was uncertain exactly how much material was released from the Hanford site, it was apparent that hundreds of thousands of Curies (Ci) of ^{131}I were released into the atmosphere around Hanford between 1944 and 1956. Table II.A-1 displays two sets of estimates of annual emissions of ^{131}I from the fuels separation processing activities between 1944 and 1957. The largest releases occurred from 1944 through 1947, and in 1949 and 1951. Radiation monitoring data recently made available indicate that these atmospheric releases of ^{131}I were carried by prevailing winds and deposited in areas surrounding the Site. During the period of largest releases, Hanford scientists gradually discovered that ^{131}I deposited readily on sagebrush and sand (4). When the soil was disturbed by wind, construction, or agriculture, this material was subsequently re-circulated and re-deposited. Thus, attempts were made to establish tolerable limits for ^{131}I on vegetation (5) and to monitor ^{131}I levels in range animals (cattle and sheep) and jackrabbits (6). Results of such studies indicate that animals were heavily exposed in areas downwind of the Site (7), and that vegetation contamination levels on-site, particularly in the 200-area, were seldom below what were considered to be tolerable at that time (.20 μCi/kg) (8). In fact, an experiment conducted in December of 1949 deliberately released a cloud of ^{131}I into the atmosphere which drifted southeastward and northeastward from the Site causing vegetation readings as high as 107.3 μCi/kg in Kennewick (9). Releases from this so-called "Green Run" have been estimated to have been as high as 7780 Ci (10) or even 11,000 Ci (10635).

Based on the data that initially became available, preliminary estimates were made of maximum doses to the thyroid that could have been received by persons living in close proximity to the Hanford Site during the years of atmospheric releases (11,12). Using environmental monitoring data for ^{131}I concentrations in vegetation, a variety of assumptions regarding agricultural production and dietary practices, and a U.S. Nuclear Regulatory Commission model to estimate thyroid doses, maximum doses were estimated (12) for residents of Richland and Pasco for four age groups: infants (0-1 yr), children (2-12 yr), teenagers (13-19 yr), and adults (20 and older). During the year of peak releases (1945), it is

estimated that the maximum annual thyroid dose to an infant may have been as high as approximately 2000 rad (11). Through 1947, maximum annual infant thyroid doses may have remained quite high (above 100 rad) with even higher periodic doses corresponding to larger atmospheric releases in the late 1940s and early 1950s. Similar estimates for infants have been proposed (11,13), with maximum annual thyroid doses decreasing to about one half these levels for children, about one quarter for teenagers, and about one-fifth for adults (9).

Table II.A-1. Estimates of Atmospheric Emissions of Radioactive Iodine from the Separations Plants Stacks

	^{131}I Ci (annual)	
Year	Anderson and Roberts Estimates*	Conklin Estimates**
1944	1700	54,000
1945	340,000	340,000
1946	76,000	76,000
1947	24,000	24,000
1948	1200	1200
1949	4670	7026 +
1950	2150	2734
1951	18,700	18,798
1952	967	996
1953	720	726
1954	540	544
1955	1200	1167
1956	370	NE++
1957	380	NE++
1944-1957	472,597	527,191

* Anderson JD. Emitted and Decayed Values of Radionuclides in Gaseous Wastes Discharged to the Atmosphere. ARH-3026, Atlantic Richfield Hanford Company, Richland Washington, 3/1/74.

* Roberts RE. History of Airborne Contamination and Control-200 Areas. HW-55569 RD, Hanford Atomic Products Operations, Richland Washington, 4/1/58, pg. 6.

** Conklin AW. *Releases of Radioactivity from Hanford, 1944-1956. Memorandum dated July 1, 1987*. Department of Social and Health Services, Office of Radiation Protection, Olympia, Washington, 1989.

\+ Includes releases from the "Green Run". Recent estimates indicate these releases may have been higher (7789 Ci) than those shown (8).

++ NE – Not Estimated.

The disclosure of information in 1986 prompted widespread concern among people living near the Hanford Nuclear Site. Residents questioned whether such releases in the past may have increased their risk for developing disease, particularly cancer. Partially in response to their concerns, a panel of experts (the Hanford Health Effects Review Panel) was convened by the Centers for Disease Control in August 1986 to evaluate the data contained in the first 19,000 pages of documents. The Panel concluded that substantial quantities of radionuclides, particularly ^{131}I, had been released between 1944 and 1956 and that off-site radiation exposures, particularly to the thyroid, were probably high enough to warrant further study of health effects. Since ^{131}I concentrates in the thyroid, it was felt reasonable to expect that potential adverse health effects associated with the Hanford radiation releases would most likely be diseases of the thyroid. Thus, the Panel recommended: 1) a study of ^{131}I releases to estimate radiation doses that could have been received by area residents, and 2) a study of thyroid morbidity among persons known or suspected to have been exposed.

A second group was also formed during this time period (March 1986), the Hanford Historical Documents Review (HHDR) Committee, which consisted of representatives from the states of Oregon and

Washington as well as several Native American tribes. A Peer Review Panel of experts was appointed as an advisory group to the committee. The HHDR focused their activities on further review of the declassified documents, and worked to consider specific approaches to implementing the two principal recommendations of the Hanford Health Effects Review Panel.

As a result of these collective efforts, a comprehensive study of potential radiation doses began in 1987, initially funded by the United States Department of Energy. The objective of the Hanford Environmental Dose Reconstruction (HEDR) Project was to develop estimates of radiation doses that people may have received from Hanford operations. A primary focus of this effort was to estimate doses to the thyroid resulting from ^{131}I exposures. Preliminary evidence from the HEDR Project indicated that the contributions to thyroid dose from the shorter-lived isotopes of iodine (^{132}I, ^{133}I, ^{135}I) were probably negligible.

Directed by an independent Technical Steering Panel (TSP) of eighteen scientists and community representatives, Battelle Pacific Northwest Laboratories in Richland performed the technical work for HEDR. While originally performed under contract to the Department of Energy, in 1993 funding responsibility was transferred to the Centers for Disease Control and Prevention.

In July 1990, the TSP made public draft reports of Phase 1 of the HEDR Project. The objective of that phase was to establish, in terms of data availability and modeling capability, the feasibility of developing a system for estimating individual radiation doses and the uncertainties of those dose estimates. For radionuclides released to the atmosphere, this was accomplished by developing preliminary estimates of doses to the thyroid from ^{131}I (14). Although external exposure (immersion and groundshine), inhalation, and vegetable consumption pathways were considered, the primary emphasis was on the cow's milk pathway for ^{131}I, since this was anticipated to be the dominant source of exposure for many people. The Phase 1 region consisted of ten counties surrounding the Hanford Site. The population of that ten-county area was approximately 270,000 during the late 1940s. For this entire population, the median dose to the thyroid from ^{131}I ingestion of contaminated cow's milk during the period 1944-1947 was estimated to be 1.7 rad, and the 90th percentile was 15 rad. It was estimated that between 1.5% and 2% of the doses for this population exceeded 100 rad.

The preliminary results from HEDR Phase 1 also identified subpopulations that received generally higher exposures. In particular, infants and young children who drank milk from family cows that grazed on pasture in areas to the east, southeast, and south of Hanford may have received substantially higher exposures. Among such children, the median and 95th percentile doses were about 70 and 650 rad, respectively. Similar children living in this area who drank commercially produced milk had a distribution of doses nearly as high.

These preliminary estimates were refined in the later phases of the HEDR Project. The total amount of ^{131}I released into the air from Hanford between 1944 and 1972 was estimated in the HEDR model to be about 740,000 Ci (2.73×10^7 gigabecquerels), with 99.8% released through 1957 (15). The HEDR results, released in April 1994, contained thyroid radiation dose estimates for representative individuals who lived in areas surrounding the Hanford facility during the times of the radiation releases, and revealed that the deposition of radioactive ^{131}I was carried further from the Site than estimated in the Phase I results. Thus, relatively less radioactive ^{131}I was deposited in areas closest to the Site, while larger amounts were deposited further away than previously anticipated. This effectively decreased the highest dose estimates, while increasing the number of people with doses in the mid- and lower ranges. Thus, while the Phase II estimates indicated lower doses than those estimated in Phase I, the results continued to provide strong evidence that large numbers of people, particularly children, may have been exposed to thyroid doses in the range of 3 to 10 rad.

One of the major products of the HEDR Project was a collection of computer programs and databases that implemented the final HEDR models for calculating doses from radionuclides released into the environment from Hanford (16). One integrated set of these programs provided estimates of thyroid

radiation doses from Hanford's atmospheric releases of ^{131}I. This included models for the amounts of ^{131}I released into the atmosphere, for the transport of that ^{131}I through the air and its deposition onto vegetation and the ground, for its uptake into food and milk products and the distribution of those products, and for the calculation of thyroid dose from exposure to ^{131}I in environmental media (air, ground, and milk and other foods). In particular, a computer program called CIDER ("Calculation of Individual Doses from Environmental Radionuclides") combined data regarding estimated concentrations of ^{131}I in environmental media with information regarding characteristics of exposed individuals (e.g., location, diet, milk and food sources) to calculate individualized estimates of thyroid dose. The HEDR Project used the CIDER model to estimate thyroid doses for hypothetical representative individuals (17,18). As described elsewhere in this report, the HTDS also used the CIDER program to calculate dose estimates for the study participants.

The second principal recommendation of the Hanford Health Effects Review Panel, the initiation of a comprehensive thyroid morbidity study, was enabled by an act of Congress in 1988. Mandated by Senate Bill 2889, the CDC was directed to conduct a study of thyroid morbidity among persons who lived near the Hanford Nuclear Site between 1944 and 1957 (Appendix 2). Thyroid diseases were selected as the primary focus for a health outcome study based on the information described above regarding radiation releases, which suggested that ^{131}I was the radionuclide most likely to pose a risk to human health. As reviewed in more detail in the section below, such exposures would be most likely to result in thyroid morbidity as opposed to other forms of illness or disability.

On March 27, 1989, the CDC issued a Request For Proposals (RFP) (Number 200-89-0716 P) to solicit applications from organizations wishing to conduct such an investigation. The proposal submitted by a team of investigators at the Fred Hutchinson Cancer Research Center and the University of Washington in Seattle was selected by the CDC, and a contract was awarded to the FHCRC on September 19, 1989.

B. Ionizing Radiation and Thyroid Disease

Radiation-induced thyroid disease in humans has generally been considered in two broad categories: thyroid neoplasia (benign and malignant neoplasms) and hypothyroidism. More recently, it has been suggested that the risk of autoimmune thyroid disease may also be increased by radiation exposure (13). In addition, acute thyroiditis can occur after high doses of radiation from orally administered ^{131}I in the treatment of certain thyroid disorders, such as thyroid cancer and hyperthyroidism (14). The degree to which the thyroid is ablated by radiation exposure, and the degree to which thyroid neoplasms or hypothyroidism result, is dependent upon several factors: type of radiation, dose, dose rate, age at exposure, sex, and current age. The type of radiation causing such disorders may be classified as either external (primarily gamma or x-radiation) or internal radiation (primarily beta) from radioiodine.

B.1. Thyroid Neoplasia: Exposure to External Photon Radiation

The evidence linking ionizing radiation with the development of thyroid neoplasms in humans has arisen largely from two sources: 1) studies of people who were previously exposed to external radiation in childhood for treatment of benign diseases of the head and neck (15-26); and 2) studies of Japanese survivors of the bombings of Hiroshima and Nagasaki who were exposed primarily to external radiation (27-28). The first category of exposures includes children treated with external radiation for acne, tonsilar hypertrophy, cervical adenitis, fungal infections of the scalp, suspected thymic enlargement (chest), and pertussis (chest). Although the first article describing the use of external radiation as therapy for such problems was published in 1907 (29), it was not until the 1950s that increased rates of thyroid neoplasia in exposed individuals began to be recognized (15-17).

Current evidence suggests that there is a dose level above which radiation-induced carcinogenesis occurs less frequently than at lower doses (30). Animal data and limited human studies collectively suggest that at external radiation doses over 20,000 milligray (mGy) or perhaps 15,000 mGy, cell killing and sterilization reduce the risk of carcinogenesis (17, 18, 25, 26). Thus, estimates of risk for thyroid neoplasia from external radiation are based on doses to the thyroid of less than 15,000 mGy.

Several cohorts exposed in childhood to external gamma radiation have been followed and evaluated for the subsequent development of thyroid neoplasia. An overall summary of such studies is difficult because important factors such as dose, age at exposure, and length of follow-up have differed. Nevertheless, these studies collectively demonstrate a dose-response relationship between external radiation dose and the development of benign thyroid adenomas and thyroid cancers (14). There have been six principal studies involving populations exposed to external radiation (20, 31-35). The range of median doses evaluated has been between approximately 60 to 8080 mGy to the thyroid. Estimates of absolute excess risk of thyroid cancer range from 0 to approximately 4 cases per million person-year-rad (PYR), averaging about 2.5 per million PYR. Among people exposed in childhood to external radiation, the absolute excess risk for total thyroid nodules has been reported to be 12.3 excess cases per million PYR (which includes thyroid cancer). A study of Israeli children irradiated for *tinea capitis* revealed higher absolute risk estimates (14 per million PYR) resulting from lower thyroid doses (average 90 mGy; range 43-169 mGy) (20).

Ron et al. analyzed the primary data from seven previously published studies of persons exposed to external radiation (36). These data showed a linear dose-response for individuals developing thyroid cancer if they were exposed before age 15. This linearity was observed down to a dose of 100 mGy but leveled at higher doses greater than 10,000 mGy. For persons exposed in childhood the excess relative risk per Gray (ERR per Gy) was 7.7 (95% confidence interval [CI] 2.1, 28.7) whereas little risk was observed for individuals exposed after age 20.

Studies of Japanese A-bomb survivors, who were exposed primarily to whole-body external radiation, show a similar dose-response relationship for thyroid cancer based on T65DR dosimetry (27).

The latest follow-up of the Japanese cohort confirms a strong dose-response for thyroid cancer. The crude incidence rates (cases per 10,000 person-years) for three dose groups (<10 millisieverts [mSv], 10-990 mSv, and >1000 mSv) showed a marked increase with increasing dose: 1.08 for the comparison group, 1.49 for the low dose group, and 3.71 for the high dose group (28). In addition, a strong linear dose-response was shown with an estimated excess RR at 1000 mSv of 1.15 (95% CI 0.48, 2.14). Age at exposure was a significant modifier of thyroid cancer risk. The excess RR at 1000 mSv was 9.46 (95% CI 4.11, 18.86) for persons exposed under age 10, compared to 3.02 for persons exposed between the ages of 10-19. These results at young ages of exposure contrasted with those for exposure after age 20, for whom the excess RR was 0.10 (95% CI -0.23, 0.75), consistent with no increased risk of thyroid cancer.

B.2. Thyroid Neoplasia: Exposure to Radioactive Iodine

B.2.a. Medical Exposures to Radioiodine

Although animal studies clearly indicate that ^{131}I can induce thyroid cancer (37-39), much less information is available in relation to the induction of thyroid neoplasia in humans from doses due to ^{131}I. Evidence from human populations arises from two principal sources: persons receiving therapeutic (moderately high) doses of ^{131}I for Graves disease or thyrotoxicosis, and persons who received diagnostic (lower) doses for thyroid nuclear ^{131}I scans to evaluate suspected thyroid disease. The early studies of persons receiving therapeutic ^{131}I for hyperthyroidism have shown no convincing evidence that the risk of thyroid cancer is increased among persons receiving ^{131}I (40-44). Most of the participants in those studies were adults at the time of exposure, were followed for very short periods, had existing thyroid disease at the time of treatment, and were treated with radiation doses that were quite high (generally 20,000 - 100,000 mGy).

A long-term follow-up of one of these studies (40) was recently published (45). This study compared cancer mortality rates in patients previously treated with ^{131}I, usually for Graves disease, to expected mortality rates for the general US population. Although no increase in total cancer mortality was found for patients treated with ^{131}I, an increase in the risk of death from thyroid cancer was demonstrated. The standardized mortality ratio (SMR) for thyroid cancer was 3.94 (95% CI 2.52, 5.86). While this increased risk was statistically significant, the absolute numbers of excess deaths were small and the authors suspected that underlying thyroid disease at the time of ^{131}I treatment might have contributed to these results.

Similar results were obtained from another recent study which evaluated cancer incidence and mortality in 7400 patients who were treated with radioiodine from 1950 to 1991 in England (#10138). The mean age of the cohort was 56 and the mean ^{131}I administered activity was 308 MBq (8.316 millicuries). The incidence and mortality rates were compared to registry data for England and Wales. Overall cancer incidence in the patient cohort was decreased (standardized incidence ratio [SIR] 0.83, 95% CI .77, .90) as was overall cancer mortality (SMR 0.90, 95% CI .82, .98). In contrast, the incidence and mortality of thyroid cancer were increased approximately 3-fold (SIR 3.25, 95% CI 1.69, 6.25 and SMR 2.78 95% CI 1.16, 6.67). However, the absolute numbers of thyroid cancer cases and deaths were quite small and the authors could not distinguish between underlying thyrotoxicosis versus radioiodine as the cause of the increased thyroid cancer incidence and mortality.

A number of studies have evaluated persons exposed to much lower doses (generally 500-1000 mGy) through diagnostic procedures (46-50). Hall et al. (49) reported in 1996 a 40-year follow-up experience of 34,000 patients who had received ^{131}I for diagnostic purposes. The mean dose for this cohort was 1100 mGy. The SIR for thyroid cancer was 1.35 (95% CI 1.05, 1.71). Excess thyroid cancers were apparent only among patients who were originally suspected of having a thyroid tumor, whereas no increased risk was noted for those referred for other reasons (49). In the group referred for suspected thyroid tumors, the increased risk was not related to thyroid dose, age at exposure, or time since exposure. The mean age at exposure of this cohort was 43; although 2408 members of the cohort (7%) were less than

age 20, stratification of risk by very young age was not reported. The authors of this study concluded that the small increase in thyroid cancer was likely due to the underlying thyroid condition and not radiation exposure. The data also suggested that protraction of dose (lower dose rate) might result in lower risk than an acute exposure of x-rays of the same total dose.

In 1989, the Food and Drug Administration's Center for Devices and Radiological Health reported the risk of thyroid disease from diagnostic ^{131}I in a cohort comprised exclusively of children and adolescents (50). Of 3483 children in the exposed group, 48% were less than 10 years and 24% were less than 5 years at time of entry into the cohort. The average length of follow-up was 27 years; the mean and median thyroid radiation dose were both less than 500 mGy. The exposed group and two separate control groups were sent questionnaires inquiring about subsequent thyroid surgery. Of 34 patients with thyroid surgery, 20 were included in the analysis. Among these 20, the proportions with malignant tumors or with benign thyroid conditions were higher in the exposed group than in either control group, however none of these differences was statistically significant (50).

B.2.b. Environmental Exposures to Radioiodine

Until 1990, the principal sources of information regarding the risk of radiation-induced thyroid disease from environmental exposures were limited to studies of Utah schoolchildren and Marshall Islanders exposed to fallout from atmospheric nuclear testing. Since then, a dramatic increase in childhood thyroid cancer has been documented from radiation exposure from the Chernobyl accident in 1986, and additional follow-up data have been published for the Utah cohort exposed to fallout from the Nevada Test Site. In contrast to the medical exposures described above, which were due exclusively to ^{131}I, most of these environmental exposures contained a mixture of ^{131}I, external radiation, and short-lived radioiodines. The following section is a brief summary of studies that have investigated the risk of thyroid neoplasia from environmental exposures to radioiodine.

B.2.b.1. Utah

Over 100 atmospheric nuclear tests were conducted at the Nevada Test Site between 1951 and 1958. Initial studies of thyroid disease incidence in Utah schoolchildren appeared to show no difference in thyroid disease outcomes compared to children from unexposed areas (51, 52). However, a follow-up study of this cohort, published by Kerber, et al. in 1993, reported an excess risk of thyroid neoplasms that was associated with exposure to radioiodine from the Nevada Test Site (53).

In that study, a relative risk of 3.4 (95% CI 0.5, 26.9) for the period prevalence of thyroid neoplasms (benign and malignant) during 1965-1986 was observed participants with estimated thyroid doses >400 mGy. A statistically significant excess relative risk of 0.7% per mGy (with 95% lower confidence bound 0.074%) was observed for total neoplasms (benign and malignant). Although positive dose-response trends were noted for total nodules and thyroid cancer (when analyzed separately), these were not statistically significant. Among 3545 study participants for whom thyroid doses could be estimated, the mean dose was 98 mGy, although for those who were children in the most heavily contaminated study county (Washington County, Utah), the mean dose was about 170 mGy. Although the dose was reported to be primarily from ^{131}I, the contribution of external radiation or short-lived radioiodines is uncertain. The authors report that the study conclusions were limited by small numbers of exposed individuals and a low incidence of thyroid neoplasms.

B.2.b.2. Marshall Islands

Of the 66 atomic tests conducted in the Marshall Islands between 1946 and 1958, the BRAVO thermonuclear test on March 1, 1954 produced the largest single radiation exposure to the Marshallese people. Extensive evaluation of this population by Brookhaven National Laboratory has shown an increase in benign and malignant thyroid nodules in residents of the northern atolls of Rongelap and Utirik (54,55).

Thyroid doses have been estimated to be primarily from a mixture of the short-lived radioiodines (^{132}I, ^{133}I, ^{135}I) and to a lesser extent, ^{131}I and external gamma radiation (56,57). For thyroid nodules, the absolute excess risk coefficient for Marshallese people from Rongelap and Utrik was reported to be 830 cases per Gy per million persons per year, or 8.3 cases per million PYR (57).

A more recent update by the Brookhaven group showed little change in prevalence of thyroid nodularity among Rongelap and Utrik residents (58). These authors also reviewed prior estimated thyroid doses in the exposed persons. For the Rongelap group, the estimated mean dose was 25,630 mGy for those with benign nodules and 16,300 mGy for those with malignant disease. For the less exposed Utrik group, the estimated mean dose was 3710 mGy for benign nodularity and 2780 mGy for malignant disease.

Although the Brookhaven studies have maintained that fallout exposure from the BRAVO test affected only the atolls Rongelap and Utrik, additional dosimetry studies have suggested a much wider area of fallout exposure (59-62). In addition, a retrospective cohort study of over 7000 Marshall Islanders showed that the prevalence of palpable thyroid nodularity (≥ 1.0 cm) decreased linearly with increased distance from the Bikini test site (63). These results were highly statistically significant and strongly suggested that fallout radiation affected a much wider region of northern and central atolls, including those with populations used by Brookhaven as controls. A new absolute risk coefficient of 1100 excess cases of thyroid nodules per Gy per million persons per year (11 cases per million PYR) was calculated using a revised estimate of the prevalence for unexposed Marshall Islanders (63). These authors also concurred with others that the exposure to the BRAVO test fallout (reported to be primarily short-lived radioiodines) appeared to be nearly as effective as external radiation in producing both benign and malignant thyroid neoplasms (54, 57, 63).

The authors of a recent report (60) attempted to independently assess the prevalence of thyroid nodularity in the Marshall Islanders, and to compare their results to the 1987 study described above. They reported a much higher prevalence of thyroid nodules in the population and a relationship between thyroid nodules prevalence and distance to Bikini atoll which was only of borderline statistical significance. However, the apparently increased prevalence can be explained in part by the inclusion of ultrasound abnormalities along with palpable nodules in their criteria for thyroid nodules. Also, since they screened very small numbers of persons from each atoll in the Marshall Islands, their study had little statistical power to detect a relationship between thyroid nodule prevalence and distance from the Bikini test site. Therefore their results cannot be viewed as inconsistent with the earlier reports.

B.2.b.3 Chernobyl

Beginning in 1992, articles began to appear reporting increased rates of thyroid cancer in children who were exposed to radiation from the Chernobyl accident in April of 1986 (64, 65). Marked increases in childhood thyroid cancer have since been reported for areas surrounding the Chernobyl reactor especially in Belarus and Ukraine (66-68). Pacini et al. evaluated thyroid cancer cases reported from registries in Belarus since 1986 and compared them with presumably unexposed cases reported from registries in France and Italy (66). Of 472 cases of thyroid cancer from six regions in Belarus, 52% were from Gomel, the most heavily contaminated region of Belarus; the numbers of cancer cases throughout Belarus roughly correlated to the degree of radioactive contamination. In addition, the Belarussian cases, when compared to the French and Italian cases, were younger and more likely to have cancers that were aggressive at initial presentation and papillary in histology. Correlations of population rates with population measures of radiation dose (e.g. collective dose) have been reported in Ukraine and Russia as well (11465). Increased rates of thyroid cancer among those who were young at exposure have also been reported in Ukraine (69) and Russia (11466).

Despite considerable efforts to assess the occurrence of thyroid cancer after the Chernobyl accident, and to determine to what extent changes in occurrence since the accident are due to radiation exposure, there is very little published information assessing a dose-response relationship between Chernobyl radiation exposure and thyroid cancer based on individual estimates of radiation dose to the

thyroid. A recent report by Astakhova et al. (11464) is probably the best attempt to date, but individual doses to children were nevertheless inferred from village Cs-137 measurements. Based on 107 cases under age 15 at the time of the accident, a strong relationship was found between estimated thyroid dose and the risk of thyroid cancer.

The radiation exposure received by people living near Chernobyl was in large part due to ^{131}I, although external radiation as well as short-lived radioiodines also contributed to the dose. Several dose reconstruction efforts have published representative thyroid dose estimates that span a wide range. Stepanenko et al. reported thyroid doses for the heavily contaminated regions of Bryansk Oblast ranging from 1600 to 2800 mGy for infants less than 1 year, and 1000 to 1800 mGy for children age 3-6 years (70). Gavrilin et al. reported estimated average thyroid doses for 14 exposed territories in Gomel and Mogilev which ranged from 220 mGy to 4700 mGy for children up to 7 years and 150 mGy to 3100 mGy in children up to 18 years (71). Likhtarev and colleagues reported estimated thyroid dose distributions in persons from five oblasts in Ukraine which showed that almost 90% of the doses in children up to age 7 were between 5 and 1000 mGy (72).

Thus, although there is now compelling evidence that the radiation exposures from Chernobyl have increased the risk of thyroid cancer in children in contaminated areas, and it is possible to estimate the range of thyroid doses received by populations in those areas, at present there is little quantitative information based on individual dose estimates regarding the risk of radiation-induced thyroid cancer after the Chernobyl accident. Furthermore, few studies have adequately addressed the potential for other factors such as iodine deficiency to modify the risk of radiation-induced thyroid cancer from Chernobyl.

B.2.c. Relative Biological Effectiveness of ^{131}I in the Induction of Thyroid Cancer

The lack of clear human evidence regarding ^{131}I induced thyroid neoplasia makes it particularly difficult to estimate the relative biological effectiveness of ^{131}I compared to external radiation in the induction of thyroid cancer. The National Council on Radiation Protection (NCRP) has reviewed data from many animal studies which have suggested that ^{131}I is from 1/2 to less than 1/20th as effective as external radiation in inducing thyroid cancer (30). One study showed that ^{131}I was equally effective to external radiation in causing thyroid cancer in Long Island rats although the effect was dependent on the presence of increased TSH stimulation (39). Based on human experience, the relative biological effectiveness was thought to be between zero and one-half. In reviewing the results of both animal and human studies, the NCRP suggested in its 1985 report that ^{131}I is one-third as effective as external radiation in producing thyroid cancer in the general population (30). It should be noted that this was intended by NCRP as a conservative value for radiation protection standards as opposed to risk estimation. The BEIR V report suggested that the radiation dose from internally deposited ^{131}I may be two-thirds as effective as external photon irradiation (73). A new NCRP report on this issue is expected but is not published at the time of this writing.

Several factors may be important in explaining a differential effect of ^{131}I as a carcinogen relative to external radiation. These factors include dose rate and the relative heterogeneity of the distribution within the thyroid gland of the dose from ^{131}I, compared to the more homogeneous dose from external radiation. Although information is limited, several studies suggest that protraction of the exposure with reduction of the dose rate may decrease the risk of developing thyroid cancer. As noted above, Hall et al. (49) suggested that the lack of radiation effect they observed in persons receiving diagnostic doses of ^{131}I may be related to the lower dose rate of ^{131}I, since the dose from a single administration of ^{131}I is delivered over a 6 week period. They speculated that this may be sufficient time for DNA repair to occur. Ron et al. also examined the effect of external radiation dose fractionation on the risk of developing thyroid neoplasia (10137). They pooled the results of three studies that included fractionated exposures and found a 30% reduction in excess relative risk (ERR) per Gy for persons whose total dose was accumulated over 2 or more exposures.

One additional study has specifically examined the effect of dose rate in children who were given external radiation for skin hemangiomas. A total of 396 children were examined at a mean of 22 years after receiving radiation in infancy (mean total dose 86 mGy) for either a short duration (seconds to a few minutes) or longer duration (30 minutes to several hours). The risk of developing a thyroid nodule increased with total dose and appeared to be linked to doses that were delivered in short duration. Although no correlation with dose was found for children exposed for only long duration, the correlation with dose for short duration approached, but did not achieve, statistical significance (ERR per Gy=10, $p<0.2$) (74). These authors suggested that dose rate may play a role in the risk of developing thyroid neoplasia from external radiation exposure.

B.3. Hypothyroidism

External ionizing radiation to the thyroid has been documented to induce hypothyroidism, although generally at high doses. Maxon reviewed a number of studies which found no clinical hypothyroidism in people who were followed up to 24 years after exposure to doses up to 10,000 mGy to the head and neck (10062). This review also included data on people receiving high doses of external radiation who developed hypothyroidism. These were typically case reports or series of patients receiving radiation therapy for malignancies such as lymphoma. Although the data are limited, the authors concluded that the induction of hypothyroidism from external radiation was likely only at doses above 10,000 mGy.

More information is available regarding the risk of hypothyroidism following radioiodine exposure. Maxon reported the risk of hypothyroidism in 6000 patients given a single dose of ^{131}I for the treatment of hyperthyroidism (10062). A strong linear dose-response between thyroid dose and the probability of hypothyroidism at five years after treatment was observed. The dose range was 25,000 mGy (minimum dose) to 200,000 mGy. The probability of hypothyroidism was 50% at five years for persons treated with 200,000 mGy of ^{131}I. These data suggested that at the minimum treatment dose of 25,000 mGy, the probability of hypothyroidism was approximately 15% at five years. It should be noted that the risk of hypothyroidism from ^{131}I in patients with Graves disease may not be generalizable to the general population.

Hypothyroidism was among the disease outcomes investigated in the Utah Study (10418). The period prevalence of hypothyroidism during 1965 through 1986 tended to decrease with increasing estimated dose. The relative risk for those with estimated doses > 400 mGy was 0.3 (95% confidence interval 0.0, 2.2), thus providing no evidence that exposure to fallout from the Nevada Test Site was associated with in increased risk of hypothyroidism.

B.4 Autoimmune Thyroiditis

Two recent studies have suggested that exposure to ionizing radiation may be associated with an increased risk of autoimmune thyroiditis. In a follow up of the Nagasaki Adult Health Study cohort of Japanese A-bomb survivors, the dose-response relationship between the prevalence of autoimmune hypothyroidism and radiation exposure was evaluated. Autoimmune hypothyroidism was defined as any TSH elevation with positive thyroid autoantibodies. Either a positive anti-microsomal antibody or positive anti-thyroglobulin antibody was considered a positive result. A dose-response was reported for antithyroid antibody positivity in persons with spontaneous hypothyroidism (13). This result suggested that exposure to external radiation might be a risk factor for developing autoimmune thyroiditis with hypothyroidism. However, the published report provided very limited information, showing only a linear-quadratic dose-response that was described as significant at the 5% critical level.

A similar result was observed in children exposed to Chernobyl fallout radiation (66). Of 171 Belarussian children, 46% had positive anti-TPO levels compared to 23% of 103 children from Italy.

Higher levels of anti-thyroglobulin were also seen in the Belarussian children compared to the Italian children. The authors postulated that thyroid autoimmune reactions may be related to radiation exposure.

Although additional data are needed to confirm an association of autoimmune thyroiditis with radiation exposure, one can speculate about potential mechanisms. One question would be whether radiation might be triggering an autoimmune response having the same natural history as spontaneous autoimmune thyroiditis with the propensity toward developing hypothyroidism. Alternatively, radiation might be causing a secondary, nonspecific autoimmune reaction resulting from damage to thyroid tissue.

C. Ionizing Radiation and Parathyroid Disease

Although the primary purpose of the HTDS was to determine whether thyroid disease is increased among persons exposed to radioactive iodine released from Hanford, a secondary objective was to determine whether persons exposed to radioactive iodine from Hanford are at an increased risk of developing hyperparathyroidism. Because the parathyroid glands are located close to the thyroid, it is possible that they may receive a radiation dose from beta-emitting ^{131}I taken up by adjacent thyroid cells. In considering potential health effects associated with thyroid radiation exposure, it may therefore be important to include effects on the parathyroid glands.

C.1. Hyperparathyroidism: Exposure to External Photon Radiation

There is considerable evidence to support the association between hyperparathyroidism and prior head and neck exposure to external beam photon radiation. Since the first case report of hyperparathyroidism in an individual exposed to head and neck radiation by Rosen, et al. in 1975 (75), there has been increasing evidence to indicate that ionizing radiation is a risk factor for the development of hyperparathyroidism.

In addition to several retrospective studies, Tisell et al. (76) reported that 14% of 444 persons who were previously treated with x-rays for tuberculous neck adenitis subsequently developed hyperparathyroidism (HPT) at least 24 years after treatment. A statistically significant dose response was found for developing HPT (dose range 0.6-45.7 Gy). For persons with doses greater than 14 Gy, 29% developed HPT with a relative risk in women twice that of men.

Cohen et al. (77) have extended their investigation of hyperparathyroidism in individuals exposed to head and neck radiation in childhood. In such persons, who had received a mean dose of approximately 8000 mGy to the tonsilar region before the age of 16, the incidence of clinical hyperparathyroidism was 18.7 per 100,000 person-years below the age of 40 and 171 per 100,000 person-years in the age range of 40 to 60 years. This represented a 2.9-fold and a 2.5-fold increase, respectively, in the incidence of hyperparathyroidism compared with that in the general population. Of interest, the above authors also found that in those persons developing hyperparathyroidism, 31% also developed thyroid cancer, compared to only 11.2% of individuals who had received prior radiation therapy but did not develop parathyroid tumors. The mean latency was 34.7 years with a maximum latency of 46 years. In addition, 90% of the cases of hyperparathyroidism were secondary to single parathyroid adenomas. In the latter study the authors recommended screening calcium measurements in the routine evaluation of persons with a prior history of childhood radiation treatments to the head and neck.

In an extension of the above study which compared prevalence rates with general population rates, Schneider et al. have more recently examined the dose-response relationship for their cohort. They report an excess relative risk of hyperparathyroidism of 0.11 per centigray in a dose range up to 100 cGy (1000 mGy) (78). The authors used dose estimates established for the thyroid; these were used as estimates of the average dose to the parathyroids.

A study of hyperparathyroidism among atomic bomb survivors in Japan corroborates the above results (79). The prevalence of hyperparathyroidism was found to be increased in individuals exposed to 500-1000 mGy when compared to unexposed control persons. A dose-response with a linear trend was observed as well as an age effect, with younger persons having higher risk.

C.2. Hyperparathyroidism: Exposure to Radioactive Iodine

Although the relationship between external beam radiation and the risk of hyperparathyroidism is reasonably well established, there is little evidence to support the existence of a relationship between radioactive iodine exposure and risk of parathyroid tumors. Animal studies have indicated that parathyroid hyperplasia or adenomas develop more frequently in rats given ^{131}I than in control animals. In addition, such studies have also suggested an age effect in rats. A higher frequency of parathyroid tumors has been observed if ^{131}I was given in the first two days of life compared to^{131}I given at 2-4 months of age (80,81).

In a retrospective report, Bondeson et al. (82) reported 600 consecutive cases of primary hyperparathyroidism of whom 10 had documented histories of prior ^{131}I treatment. Such treatment had been given for either Graves Disease or for ablation of thyroid remnants. Age at the time of ^{131}I therapy ranged from 21 to 72 years with the interval to detection of hypercalcemia ranging between 3 and 27 years. These authors also indicate that parathyroid adenomas developed at the sites of thyroid remnants in cases with ^{131}I ablation after thyroid tumor operations.

While the mechanism of parathyroid tumor induction in individuals exposed to external beam radiation is almost certainly due to direct photon beam exposure, the mechanism of postulated parathyroid tumor induction from radioactive iodine is less certain. The parathyroid glands are not known to take up iodine. However it is plausible that parathyroid cells can be exposed to beta radiation from ^{131}I taken up in thyroid cells adjacent to the parathyroid glands. This mechanism of exposure is consistent with the results summarized above since the parathyroid glands in rats are imbedded within the thyroid tissue whereas in humans they exist as separate organs. Although the number of cases is quite small in the study by Bondeson et al. (82), the development of parathyroid adenomas near the site of thyroid remnants treated with ^{131}I supports this hypothesis.

Estimated doses to the parathyroid glands can be calculated if the thyroid dose from radioactive iodine is known. For example, a 5.0 mCi administration of ^{131}I would be expected to give a thyroid dose of approximately 45,000 mGy and a parathyroid dose of approximately 16,500 mGy (83). Therefore, the parathyroid dose from ^{131}I is approximately 30% of the thyroid dose for a given amount of ^{131}I.

Thus, while it seems clear that external radiation is a risk factor for the development of parathyroid tumors and subsequent hyperparathyroidism, the association of parathyroid disease with radioactive iodine exposure is less certain. Nevertheless, the available data are suggestive and warrant further investigation.

D. Ultrasound-Detected Abnormalities of the Thyroid (Thyroid UDAs)

Since the mid-1980s, high-frequency ultrasound has increasingly been used in the evaluation of thyroid nodules. Although the traditional definition of a thyroid "nodule" is based on clinical palpation, the greater sensitivity of ultrasonography has led to its greater use, since it can detect nonpalpable, millimeter size abnormalities. Several important issues, however, have arisen with the use of this technology: 1) thyroid UDAs have been shown to occur frequently in the general population without good understanding of their risk of malignancy or biologic significance; 2) thyroid UDAs have often been classified as "nodules" regardless of size; 3) the use of ultrasound in defining criteria for thyroid nodules has made it difficult to compare clinical thyroid outcomes among epidemiological studies using different criteria for thyroid nodularity and; 4) although ultrasound has exceptional sensitivity, recent data regarding specificity (the ability to distinguish benign from malignant nodules) suggest that the increased specificity of ultrasonography is associated with a significant decrease in sensitivity.

As described further below, a few published studies have examined the possibility of association between radiation exposure and thyroid UDAs. However, to interpret those studies properly, careful attention must be paid to the issues mentioned above. The following section summarizes the published literature regarding the prevalence, clinical significance, and possible radiogenesis of thyroid UDAs.

D.1 Prevalence of Thyroid UDAs

A number of studies have shown a high prevalence of thyroid UDAs in the general population. Tan et al. recently reviewed the literature and reported a range of prevalence of 17-67% (10269). In 1000 persons referred for evaluation of hypercalcemia (of whom 8% had a nodular goiter), 46% had discrete thyroid lesions on ultrasound and 38% were reported to have thyroid nodules (10446). While these patients are unlikely to be representative of the general population, they were not referred for suspicion of thyroid disease. The highest prevalence of thyroid UDAs was reported in a prospective study of 100 female employees responding to a notice on a bulletin board: 67% of these women, mean age 43, had abnormal thyroid ultrasound scans (10114). The results of this study are limited by small numbers. Thyroid UDAs in populations without apparent thyroid disease have also been documented outside the US with prevalence figures ranging 17-27% (10777, 10229, 10444). Most of these studies have been consistent in showing that nonpalpable thyroid UDAs are generally small and that solitary nodules on clinical examination are often associated with multiple other thyroid UDAs. Both Tan (10831) and Brander (10777) have demonstrated that 48% of patients with known palpable thyroid nodules greater than 1 cm harbored additional thyroid nodules found on ultrasound.

Brander and colleagues have published two important studies. In the first study, 253 persons randomly selected from a Finnish city council registry were screened for thyroid UDAs (10683). The sample was distributed evenly among four age brackets from 20 through 50. The community was not thought to have endemic goiter. Thyroid UDAs were detected in 69 persons (27%). These abnormalities were solitary in 57%, multiple in 22%, and diffuse in 22%. The mean age for persons with normal ultrasound scans was 35, the mean age for the group with abnormal ultrasound findings was 37. The frequency of these abnormalities was higher in women than men and increased with age for both sexes. For women, the prevalence of thyroid UDAs was 30% in the 20-29 age group, 32% in the 30-39 age group, and 41% in the 40-50 age group. All participants underwent thyroid palpation prior to ultrasound examination. Palpable abnormalities were detected in 13 persons (5.1%): three with a solitary nodule, five with multiple nodules, and five with abnormal consistency. Fine needle aspirations were done in 30 individuals. All were negative for malignancy with one intermediate probability of neoplasm; that person underwent surgery and had a follicular adenoma. The authors commented that thyroid UDAs were common in an unselected population, and that the likelihood of malignancy was low. They recommended a conservative approach to these lesions.

In the second study (10209), Brander and colleagues performed follow-up ultrasound scans in persons who initially had thyroid UDAs in the previous study. Of the 69 persons with initial thyroid UDAs, 57 (83%) were located and re-evaluated 5 years later. Of these 57 persons, 28 had thyroid UDAs that were defined as macrofollicles (lesions less than or equal to 5mm). After 5 years, 14 macrofollicles were unchanged, 5 increased in size, 5 decreased in size, and 4 had no follow-up.

The remaining 29 persons had a total of 34 nodules which had been detected during the initial ultrasound screening. Of these, 12 had grown over 5 years, 8 had either disappeared or diminished in size, and in seven persons a new lesion developed. Of the 12 that had grown, biopsy was performed in 10 with 9 benign results and 1 which was a benign adenomatous nodule after surgical excision. Of the 7 new lesions, biopsy was performed in 5 and all were benign. At the end of the 5-year follow-up, there were no individuals with thyroid cancer who previously had thyroid UDAs at the initial screening. The authors acknowledged the small size of their study but concluded that in contrast to persons with nodules selected for surgery, "most lesions randomly detected at ultrasound of the thyroid are benign."

Bruneton evaluated 1000 healthy volunteers without history of thyroid disease and performed high frequency thyroid ultrasound examinations (10517). Although selection criteria or mean age were not provided, 57% of participants were over 50 years. Ultrasonography was performed with 13 MHz transducers and all ultrasound nodules greater or equal to 3 mm were counted. One or more nodules were detected in 34.7% of participants. For persons less than age 50 (n=431), the prevalence was 25%. For persons greater than age 50 (n=569), prevalence was 42%. For all ages, the prevalence in women was 44% and the prevalence in men was 17.7%.

A Belgian study assessed thyroid UDAs in 300 patients who were referred for abdominal ultrasound examinations (10444). Although this study sample is not a random representation of the general population, there were extensive exclusion criteria for those with symptoms or signs of thyroid disease. The mean age was 47 (1-88) and 55% of participants were males. Small echoic nodules were found in 19% of patients. In patients in their 7th decade, the prevalence increased to over 40%.

These ultrasound prevalence studies can be compared to the autopsy study by Mortensen in 1955 which showed that approximately 50% of 1000 consecutive autopsies had single or multiple thyroid nodules in glands which appeared "clinically normal" (10046).

D.2. Specificity of Thyroid Ultrasonography in Predicting Thyroid Cancer

There has been significant controversy regarding whether there are ultrasound characteristics that can independently predict malignancy in thyroid lesions. Rago and colleagues assessed 104 consecutive patients by conventional ultrasound and color flow doppler prior to thyroid surgery (10240). The characteristics of the halo sign, hypoechogenicity, and microcalcifications were assessed by conventional ultrasound while Type I, II, and III color flow patterns were assessed by color doppler. The combination of absent halo, the presence of microcalcifications, and a Type III color flow pattern increased specificity for thyroid cancer to 97%. However, the sensitivity decreased to only 16%. Thus, while ultrasound and color flow doppler increased the specificity for thyroid cancer it did so at the expense of sensitivity for predicting thyroid cancer.

Takashima studied the sonographic and pathologic correlation in 69 of 99 surgically removed nodules (10020). Microcalcification showed the highest specificity of 93% with a positive predictive value of 70% for thyroid cancer. However, the sensitivity was only 36%. They discussed the distinction between dense calcifications, which are found in both thyroid cancer and benign lesions, and microcalcifications which are much more specific for thyroid cancer. However these are not always seen on ultrasound but may be found on pathology review. The authors conclude that "none of the various sonographic features, such as multiplicity of nodules, presence or absence of halo or cystic areas, lesion echogenicity, or margin

characteristics help to reliably distinguish between benign and malignant thyroid nodules." They state that "microcalcifications were useful however sonographic microcalcification is not a sensitive nor sufficiently accurate indicator of malignancy because pathologic microcalcifications are found, at most, in only 60% of thyroid cancers."

Tominori evaluated the combination of ultrasonographic and cytologic characteristics in predicting thyroid cancer and developed an index which prepared patients better for selection for thyroid surgery. He acknowledged that "clearly sonographic features alone do not reliably separate benign from malignant thyroid nodules" (10229).

In a similar statement, Sakaguchi reported that studies indicate that several ultrasound characteristics are "suggestive" of thyroid cancer such as solid and hypoechoic lesions, irregular margin, and fine microcalcifications (10231). However, the authors stated that, "There is no single sonographic criterion that distinguishes benign from malignant thyroid nodules." In a recent commentary, Hegedus and Karstrup state, "A general finding – has been that there is no US [ultrasound] pattern, alone or in combination with other techniques, that may be considered specific for thyroid cancer" (10218).

D.3. Ionizing Radiation and Thyroid UDAs

The increased sensitivity and the development of portable ultrasound equipment have made ultrasonography particularly attractive in evaluating abnormalities of the thyroid gland in persons exposed to environmental radiation. In contrast to the increasing volume of literature regarding thyroid UDAs in the general population, much less is known about whether ionizing radiation causes an increase in thyroid UDAs prior to the development of clinical disease.

Schneider and coworkers evaluated a subgroup of their Michael Reese cohort who had been exposed to head and neck radiation therapy during childhood for benign conditions. They selected 54 individuals who had previously had normal thyroid exams and normal thyroid nuclear scans in the 1974-76 time period. Of these 54 persons in this follow-up study many years after exposure, 47 (87%) had one or more discrete thyroid UDAs (10111). In this cohort, external radiation exposure was clearly associated with increased thyroid UDAs. The authors concluded that: 1) thyroid nodules continued to develop in radiation-exposed individuals many years after exposure and 2) although thyroid UDAs were quite common in the general population, they were more prevalent in radiation-exposed populations.

Other studies have also suggested that thyroid UDAs are more common in exposed populations. Antonelli, et al. compared ultrasound scans of two groups: 50 hospital workers with occupational radiation exposure (external radiation) in a hospital setting and 100 controls without such exposure (10154). Thyroid UDAs were detected in 38% of the exposed persons and only 13% of the controls. Similarly, Sugenoya and colleagues (11126) compared 299 children who were exposed to Chernobyl radiation to 323 children who were unexposed. Although none of the children in either group had palpable abnormalities, 34 of the exposed (11.4%) had thyroid UDAs compared to 4 unexposed children (1.2%).

There is very limited information regarding the dose-response relationship between radiation exposure to the thyroid and thyroid UDAs. While such abnormalities might be expected to correlate with clinical thyroid disease, the question of whether thyroid UDAs might represent an early marker of radiation injury prior to the development of clinical disease is unknown. There are currently no studies in the literature to answer this question.

III. STUDY OBJECTIVES

The primary objective of the HTDS was to determine whether thyroid morbidity (including, but not limited to hypothyroidism, benign neoplasia, and malignant neoplasia) is increased among persons exposed to atmospheric releases of radioactive iodine from the Hanford Nuclear Site between 1944 and 1957. If an effect was detected, the study was designed to further determine in what way the increase in thyroid morbidity is related to the dose of radiation received (i.e., the characteristics of any dose-response relationship).

In addition to these primary objectives, the HTDS had three specific secondary objectives: 1) to determine whether hyperparathyroidism is increased among persons exposed to the Hanford radiation releases and who received radiation doses to the thyroid and, if so, to determine in what way the increase in hyperparathyroidism is related to the dose of radiation received; 2) to provide information to residents of the communities surrounding the Hanford Site regarding the objectives, design, and conduct of the study, as well as the findings and results of the research; and 3) to assess the appropriateness of the methods employed and the degree to which such an investigation could be successfully planned and executed, given the long interval since exposure and the uncertainties regarding radiation dose.

IV. STUDY DESIGN

A. Eligibility Criteria

The HTDS was based on a cohort of people defined by the following eligibility criteria:

- Mother's residence at the time of the participant's birth: Benton, Franklin, Walla Walla, Okanogan, Ferry, Stevens, or Adams County in Washington State
- Year of birth: 1940 – 1946.

The rationale for this choice of counties and years is described in sections IV.A.1 and IV.A.2 below. The mother's usual residence at the time of the participant's birth, which can be determined from birth records, was used as a criterion since it was likely to indicate the participant's place of residence during the first years of Hanford's operations, when the largest releases of ^{131}I occurred (see section V.A.2 below). The cohort included the majority of the possible combinations of the seven counties and seven birth years. However, birth year subcohorts for certain counties were not included since they were unlikely to include many participants with relatively high thyroid radiation doses (see sections V.A.2 and V.A.3 below).

A.1 Mother's Residence at the Time of the Participant's Birth

Geographical proximity to the Hanford Nuclear Site is clearly a determinant of radiation doses received by area residents. The atmospheric transport and deposition of radioactive materials depend on the location of the source of the release, the surrounding topography, and meteorological conditions at the time of the release (i.e., wind speed, wind direction, precipitation and atmospheric stability). The HEDR Project considered such factors to estimate the atmospheric dispersion of radioactive iodine from Hanford. Preliminary HEDR results were used to define the geographical boundaries for selection of the HTDS Pilot Study Sample, and final HEDR estimates used to refine the boundaries for the selection for the Transition and Full Study Samples. The 75,000 square mile geographical domain within which the final HEDR model applies is shown in Figure II.A-1 above.

The prevailing winds in the vicinity of the Hanford Site blow primarily from the North, Northwest, and West across the Site to the East and Southeast. Although there were some seasonal variations according to month of the year, this pattern was generally consistent throughout the year during the 1940s (92). Wind direction determines the directions in which airborne plumes of radioactive material most likely traveled, and thus the geographical areas most likely to have received deposits of radionuclides. For the most part, atmospheric releases traveled to areas East and Southeast of the Hanford Site.

Utilizing meteorological data from the 1980s, information regarding the amount of material released, and limited off-site monitoring data, the HEDR Project calculated ^{131}I concentrations in vegetation surrounding the Site. Later calculations conducted by Battelle Pacific Northwest Laboratories using 1944-1947 meteorological data generally confirmed this geographical pattern of ^{131}I concentration in sagebrush. These data suggested that the areas of highest concentration were primarily those closest to the Site to the East and Southeast (e.g., in Benton, Franklin, and Walla Walla counties). Final estimates from the HEDR Project revealed a wider dispersion of ^{131}I, with decreased concentrations in the areas nearest the site and increased concentrations in areas somewhat more removed especially northeast of Hanford. This finding prompted the inclusion of Adams County in the Transition and Full Study Samples.

An important pathway of radioactive iodine exposure in humans is the ingestion of milk produced by animals grazing on radioactive iodine-contaminated vegetation. Therefore, milk distribution routes and

milksheds are also important in defining a geographic area in which people were exposed to radioactive iodine. In an effort to describe such components of a milk pathway for radioactive iodine, the HEDR Project attempted to reconstruct the following types of information for the period of highest releases (94): 1) types, quantities, and sources of feed for dairy cows in the area; and 2) location, relative size, and distribution routes of all fresh milk processors/distributors in the area.

For purposes of planning the Pilot Study, it was important to know the sources and distribution patterns of the milk consumed in the areas of interest during the time period under study. The HEDR Project developed a summary measure for each of the ten counties surrounding the Site, based upon production and consumption data, to indicate whether the county recorded a milk surplus, a milk deficit, or was in relative balance regarding milk production and consumption.

Several important findings were reported (94). Overall, the 10-county area surrounding Hanford was self-sufficient in milk production and, in fact, recorded a surplus of almost 20% in 1945. There is considerable variability by county, however, which is largely explained by different amounts of irrigated pasture available for raising dairy herds. For example, Yakima, Kittitas, and Klickitat counties were milk surplus counties, particularly Yakima and Kittitas. Essentially all of the milk consumed by residents of these counties was produced locally. Similarly, all of the milk consumed in the Walla Walla area was locally produced. In contrast, Benton and Franklin counties imported milk. Two of the primary sources of commercial milk for residents of these counties were the Carnation Dairy in Sunnyside and the Twin City Creamery in Kennewick. Although the Twin City Creamery itself was located in Benton County, it received milk from a number of dairies outside the county.

In addition, two special circumstances with regard to milk supply needed to be considered in selecting potential study participants. First, much of the commercial milk consumed by Richland residents is thought to have come from the Carnation Plant in Sunnyside. In fact, the Atomic Energy Commission had a contract with Carnation in Sunnyside to supply the town of Richland with their milk during this time period (95). Second, a substantial amount of the milk consumed by people living in rural areas (which constituted much of the area during the 1940s) was supplied by backyard cows. It is estimated that between 40% and 90% of the milk consumed by rural families came from this source.

Thus, based upon these preliminary findings regarding meteorological conditions, the deposition and concentration of radioactive iodine in vegetation, and the patterns of milk production and consumption by county, the area encompassed by Benton, Franklin, and Walla Walla counties was defined in the Pilot Study as the area within which people with the highest thyroid doses were most likely to be identified. Adams County replaced Walla Walla County in the Full Study selections, as noted in section V.A.3.b, due to the findings of the HEDR Project's final report.

On July 12, 1990, the HEDR Project released its preliminary or 'Phase I' estimates of thyroid doses potentially received by residents in a 10-county area surrounding the Hanford Site (12). These estimates were based upon computer models that included information regarding the amount of radioactive material released, the dispersion and deposition of the material, the uptake of the material into the food chain, the geographical distribution of food products (especially milk and dairy products), and the consumption of contaminated food products by humans. These preliminary estimates confirmed that people who lived closest to the Hanford Site, particularly to the East and Southeast, were likely to have received the highest radiation doses to the thyroid.

To determine whether thyroid disease is increased among people who received a radiation dose to the thyroid from Hanford radioactive iodine releases, it is necessary to compare rates of thyroid disease among people with different levels of exposure, including very low exposures or no exposure. The considerations described above guided the definition of a group likely to have received the highest doses. However, the selection of groups likely to have little or no exposure, or intermediate levels of exposure, is equally important in properly evaluating whether the radiation exposure is associated with an increase in the risk of thyroid disease. The most important considerations in defining these groups are: 1) that the

groups are comparable to the high-dose group with respect to other factors which could confound any relationship between radioactive iodine exposure and thyroid disease (e.g., geography, urban/rural composition, occupational factors, socioeconomic factors, age, ethnicity, sex); and 2) that the same opportunities and resources exist to identify and trace people in low- and intermediate-dose groups as in the high-dose group.

With these requirements in mind, three types of comparisons can be considered for use in a cohort study of this type. First, rates of thyroid disease in a general population (e.g., the entire United States) could be used to evaluate whether the rates of thyroid disease observed in the cohort receiving a dose are higher than would be expected based upon general population experience. In the present context, however, this approach is problematic primarily because, with the exception of thyroid cancer, incidence rates for the thyroid diseases under study are generally not available for other populations (e.g., the United States or Europe). Estimates could be obtained from other study cohorts that have been followed subsequent to exposure to ionizing radiation (e.g., the New York tinea capitis study), but such rates do not reflect general population experience and the degree to which they can or should be generalized to the eastern Washington experience is questionable.

A second approach would be to include people who received zero or very low thyroid dose in the cohort selected from the geographical areas most likely exposed to the Hanford releases as a comparison group, and to compare disease rates in people with higher dose levels to the rates among those in this "baseline" category. An "internal" comparison group such as this has a number of advantages. Most importantly, concerns regarding comparability of other factors that might influence the risk of thyroid disease are largely resolved. Second, it is more efficient to enroll and study a single cohort, rather than two geographically separate sub-cohorts (one from areas most likely exposed and one from areas most likely unexposed). Third, this approach allows for a very flexible analysis based on both a simple dichotomy of exposure (dose vs. zero or very low dose) as well as a quantitative estimate of dose (i.e., a dose-response). For this approach to succeed, however, relatively accurate individual doses need to be available, a full range of doses needs to be represented in the cohort (or at least a sufficient number of people with relatively high doses), and an adequate number of people with zero dose need to be included.

In this study, particularly for its pilot phase, it could not be assured that the dose distribution in the sub-cohort selected from Benton, Franklin, and Walla Walla counties would allow for such an approach (e.g., that there would be an adequate number of people with very low doses, as well as a full range of higher doses) because the selection of potential study participants could not be based on estimates of an individual's radiation dose. In fact, it could not even be assured at the time the Pilot Study was initiated (1992) that an adequate individual-level dose estimation system would be available from the HEDR Project for this study. Finally, among people born in Benton, Franklin, and Walla Walla counties during 1942-1946, those who received very low doses could not be presumed to be comparable, with respect to their natural risk of thyroid disease, to those who received higher doses. Those with little or no dose would likely have drunk little or no milk, moved to locations more distant from Hanford before accumulating more dose, and drunk milk imported from a less-exposed area. These characteristics may be related to health or socioeconomic status, and while none of them is known to influence the risk of thyroid disease, the possibility that they may could not be ignored.

Consequently, a third approach for identifying comparison participants was to select people who were not likely to have been exposed to the Hanford radiation releases and who were, therefore, not likely to have received a radiation dose to the thyroid, on the basis of geographic proximity to the Hanford Site. The primary purpose of such an approach was to be able to identify potential study participants with a high degree of certainty that they received a very low or no radiation dose from radioactive iodine from Hanford. The two principal concerns with this approach, however, were: 1) to establish that such people were truly unexposed (and, therefore, received very low or zero dose) and 2) to assure that such people are comparable to those who did receive a dose regarding other characteristics that might influence the risk of developing thyroid disease.

Thus, for the Pilot Study, some of the participants were identified from areas that are geographically removed from the three-county area considered most likely exposed. This design served two important purposes. First, it enabled the evaluation of radiation doses to the thyroid for a group of people anticipated to be relatively unexposed to the atmospheric releases of radioactive iodine from the Hanford Site in a manner identical to that used for people anticipated to be highly exposed. Second, this approach provided information regarding the success rates of the different data collection aspects of the project among people born in areas removed from the Hanford Site relative to people born in areas in close proximity. It was recognized that, to the extent substantial differences in the success of participant enrollment and data collection efforts were identified, revisions in the overall study design would have been necessary.

For a separate geographical area to be suitable as a source of study participants, two conditions must be met: 1) the area must have received little or no exposure to the radioactive iodine from Hanford either directly through atmospheric transport of the iodine or from importation of agricultural products (principally milk) from contaminated areas; and 2) the participants from that area must be as comparable as possible to those from the more exposed area regarding other factors that might influence the risk of thyroid disease. In practice, it may be difficult to define an area that satisfactorily meets the conditions specified by both of these criteria. With increasing distance from the Hanford Site (and, therefore, less likelihood of exposure), there was concern that comparability of other factors would be more difficult to achieve.

Regarding exposure to the atmospheric releases of ^{131}I, initial information available from the HEDR Project about prevailing wind patterns, ^{131}I deposition, and commercial milk distribution suggested that counties to the west and northwest of the Hanford Site might be possible sources of study participants with relatively low doses. Sagebrush concentrations of ^{131}I in the northwestern-most sections of the 10-County HEDR Phase I region were two or three orders of magnitude less than those in areas immediately surrounding the Hanford Site, and these same counties were generally milk surplus areas, meaning that they received little or no milk from cattle in areas likely to be more heavily contaminated by ^{131}I.

Nevertheless, a careful review of the preliminary HEDR Phase I thyroid dose estimates released in July, 1990, indicated that even in these counties some people may have received considerable radiation doses to the thyroid; for example, infants in the eastern Census Divisions of Kittitas County (KI1, KI7, KI8, KI9, KI10, KI11) (see Figure 2.5 in reference (12)). Furthermore, as continuing efforts were made by the HEDR Project to improve the quality and completeness of the basic data regarding meteorology and milk distribution used to calculate radiation doses, it became apparent that appreciable doses possibly occurred in some of the counties outside the 10-County Phase I region (Figure IV.A-1, used with permission, 93).

Thus, based upon the information available at the time about possible thyroid doses, it seemed most prudent to attempt to locate subjects with little or no dose from areas at least one additional county "layer" distant from the Phase I boundary, and more directly to the north of the Hanford Site. Consequently, in the Pilot Study, selection of cohort members was extended to the three counties most directly north of the Hanford Site (Okanogan, Ferry, and Stevens) which are separated from the Phase I 10-county HEDR boundary by one "layer" of counties.

The criterion of comparability of factors other than radiation exposure that might cause thyroid disease was also a concern. Relatively little is well established regarding the causes of thyroid disease. Factors known to be of most potential concern were: 1) selected demographic factors, most notably age, sex, and race; 2) socioeconomic status, most importantly as it relates to access to medical and dental care and resulting exposures to medical and dental sources of ionizing radiation; and 3) dietary iodine intake.

In selecting Okanogan, Ferry, and Stevens counties for inclusion in the Pilot Study, it was important to assess the degree to which the populations of these counties were similar to those of Benton, Franklin, and Walla Walla counties in the 1940s regarding at least the factors listed above. Age and sex distributions for 13 counties in central and eastern Washington and north-central Oregon that might have

been considered sources of people exposed to little or no radioactive iodine from Hanford showed little variability among the counties in the ratio of males to females. Okanogan, Ferry, and Stevens counties were shown to have slightly higher proportions of children under the age of five than do Benton, Franklin, and Walla Walla, but the difference was relatively small.

Figure IV.A-1. Iodine-131 Thyroid Dose from All Exposure Pathways (Milk Cows on Fresh Pasture)

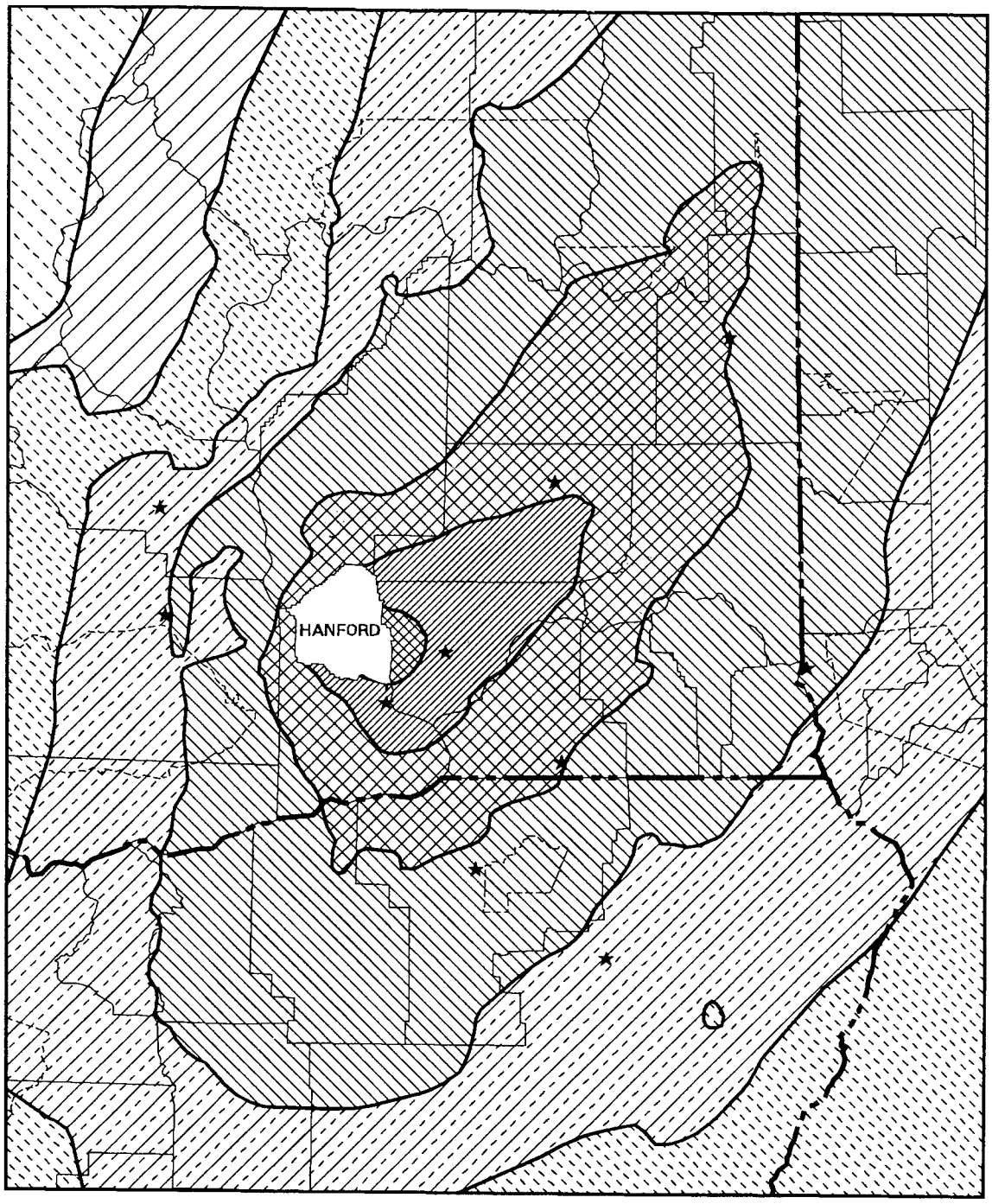


Figure IV.A-1. Legend

Comparisons of data from the 1940 Census showed that the populations of the counties to be included in the Pilot Study were overwhelmingly classified as rural and white. The median number of years of school completed, and the proportion of the population 14 years of age and older employed in five major occupational groups were also similar. Thus, it was concluded that the composition of the six counties included in the Pilot Study in terms of age, sex, race, education, percent rural, and major occupational category were reasonably similar, and not greatly different from other counties in the larger surrounding region.

It was also important to assess the degree to which iodine availability and/or intake might vary among study counties. Geographical differences in the distribution of iodine intake could result in geographic differences in the rates of one or more of the thyroid diseases under study (e.g., endemic goiter belts). To the extent that such differences might be related to radiation dose from Hanford, they could potentially confound an association between radiation exposure and thyroid disease.

Preferable to estimates of soil iodine concentrations would be estimates of iodine intake. Although little work had been conducted in this regard on a geographic basis, in 1970 Oddie et al. (96) reported estimates of average dietary iodine intake derived from thyroidal radioiodine uptakes in approximately 30,000 euthyroid subjects in 133 locations throughout the United States. Although average daily iodine intake varied considerably throughout the United States (from 240 to 740 micrograms per day), the Pacific Northwest was relatively uniform in the distribution of daily intake estimates. Mean values were reported for fifteen areas in the Northwest centered by two degrees latitude and longitude (approximately 140 by 120 miles). All values in the six Pilot Study counties were between 345 and 379 micrograms per day (a very narrow range compared to the overall distribution of values). Thus, within the confines of most of central and eastern Washington and north central Oregon, there is some evidence to suggest that iodine intake was adequate and relatively uniform in the past.

The inclusion of Okanogan, Ferry, and Stevens counties in the Pilot Study was intended to provide a convenient mechanism for identifying an adequate number of potential study participants who received little or no radiation dose to the thyroid from Hanford. It was not intended to serve as a means of defining a "comparison area" or "control group." Although potential study participants were selected based, approximately, on the county in which they were born (see section V.A., below), the fact that a person was born in one area or another was not relied upon to determine whether he or she was actually exposed to radioactive iodine from Hanford and, more importantly, actually received a radiation dose to the thyroid. Exposure, and the estimate of the resulting radiation dose to the thyroid, was determined from a detailed residential history and exposure information collected whenever possible from the mother or other close relative of each study participant (discussed more fully in section V.D. below). For example, a person born in Benton County between 1942 and 1944 may have moved to a residence away from Hanford before any exposure could occur. Similarly, a person born in Stevens county may have lived or visited for a prolonged time (e.g., a summer) within the "exposed area" and received a substantial thyroid dose.

Nevertheless, it was assumed that most of the study participants with the highest thyroid doses would come from Benton, Franklin, or Walla Walla County, and that most of the participants from Okanogan, Ferry, or Stevens counties would have very low thyroid doses. The use of separate geographic areas was simply a device that would allow a degree of control over participant selection to assure adequate numbers of participants in the Pilot Study who would have thyroid doses at the highest and lowest extremes of the dose distribution.

Based on the results from the HTDS Pilot Study, it was determined that the inclusion of geographically removed populations in the selections for the Transition and Full Study samples was unnecessary. In fact, maximizing the number of participants with the highest doses proved to be of much greater concern. Thus, no additional selections were made from Stevens, Ferry, and Okanogan counties following the Pilot Study selection. In addition, due to the HEDR Project's findings that people in Adams County could be expected to have received higher doses than those in Walla Walla County, the HTDS cohort was completed by selecting people from the Richland, Pasco/Kennewick, Benton County, Franklin

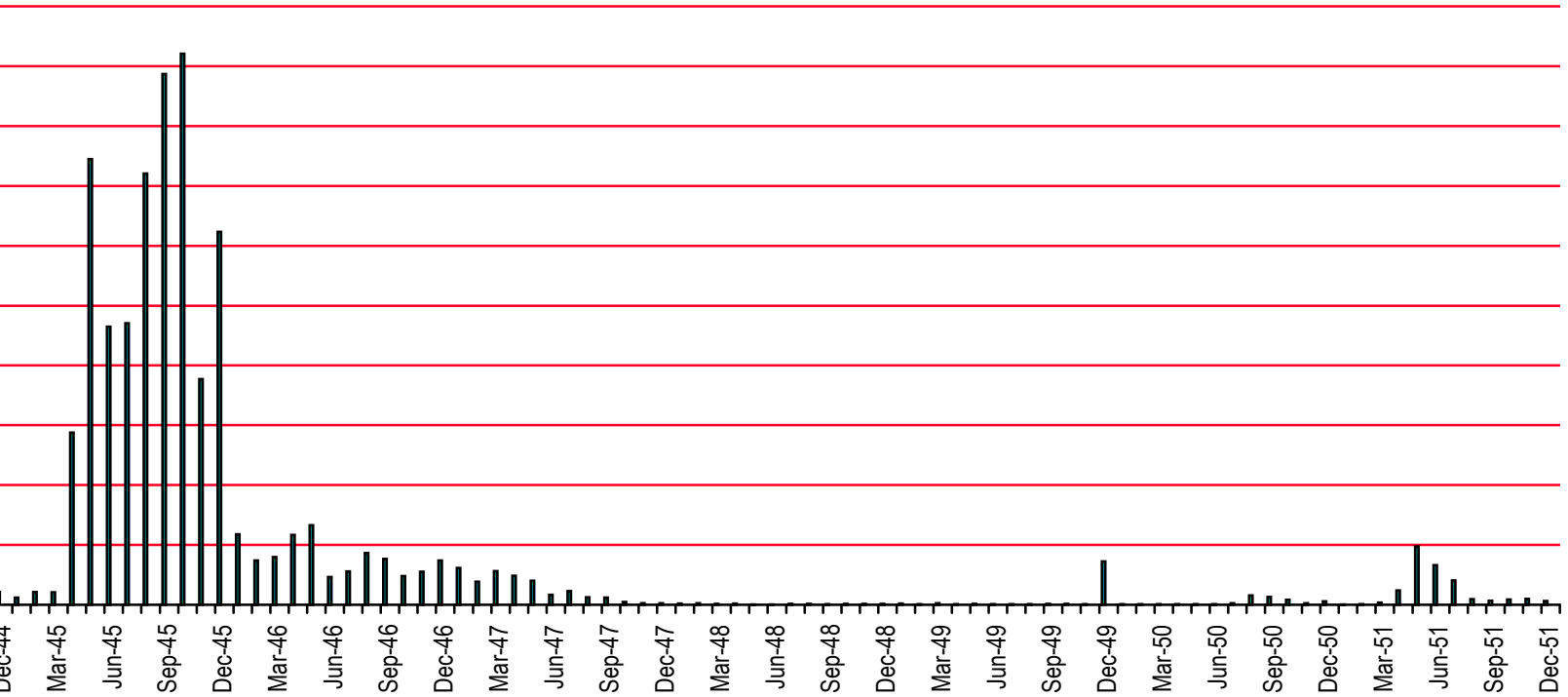
County, and Adams County geostrata.

If it was learned during any part of the study that a potential participant was adopted, the potential participant was considered eligible if verification could be obtained that the birth mother's place of residence at the time of the potential participant's birth was within one of the seven counties included in the study.

A.2. Year of Birth

The radioactivity of ^{131}I decays exponentially with a half-life of 8.4 days. This implies that nearly all of the thyroid dose produced by ^{131}I released into the environment will be accumulated within a few months after its release. Therefore the time period of most interest for identifying potential participants who could have received the highest thyroid doses is that which corresponds to the largest atmospheric releases of radioactive iodine from the Hanford facility. As shown in Figure IV.A-2, the large majority of the releases occurred from the last two weeks of 1944 through 1946. Beginning in 1947, monthly releases were considerably lower, averaging between 100 and 2000 Ci (11). The exceptions to this pattern were the substantial release associated with the "Green Run" in December 1949 and the releases during 1951. Thus, the time during which area residents would likely have received the highest exposure to radioactive iodine would have been the years 1944-1946.

Figure IV.A-2. Monthly ^{131}I Releases from the Hanford Nuclear Site, 1944-1951



Existing literature suggests that age at exposure is an important factor in radiation-induced thyroid disease. In particular, risks among those who are very young at exposure may be higher than for those who are adults at exposure for one or both of the following reasons: 1) higher radiation doses to the thyroid result per unit exposure; or 2) there is an increased sensitivity in the young (i.e., an increased risk per unit dose in the young). It is known that iodine is metabolized differently in children than in adults. The concentration of radioactive iodine in the smaller thyroids of children (the infant thyroid is only about 1/10 the size of an adult thyroid) is greater per unit exposure (30). Book (97) has shown that the thyroid dose to infants resulting from the inhalation of a fixed concentration of radioactive iodine in air is twice that of adults. The dose in near-term infants is ten times that of adults. By the ingestion pathway, six-month-olds could receive thirty times the dose of an adult, largely as the result of smaller thyroids and a higher intake of milk (30).

A number of epidemiological studies have given rise to more indirect evidence regarding the issue of increased sensitivity in the young. Dobyns et al. (40) reported an increased risk of thyroid adenoma among the youngest quartile of a cohort treated with ^{131}I for Graves disease, although risks by specific years of age were not investigated. The incidence of thyroid cancer among atomic bomb survivors in Japan exposed primarily to external gamma radiation has been shown to be higher among those exposed at young ages. A strong dose-response was seen in this cohort, with a three-fold increase in the excess relative risk of children exposed less than 10 years of age compared to those exposed at ages 10-19. Marshall Islanders exposed to nuclear fallout (external gamma and radioactive iodine) had increased rates of thyroid neoplasia at earlier ages of exposure. Compared to people exposed at age 18 and over, those exposed under age 18 had 2.5 times the risk of developing benign nodules and those exposed in utero had a five-fold risk of developing a benign thyroid nodule (55). Results from a study of people treated with radiation for tinea capitis in Israel (external gamma exposure) indicate that children exposed under the age of five had 3.1 times the number of excess thyroid cancers at age 40 than those children exposed over age five (51). Although accurate dosimetry has hampered risk assessment of thyroid cancer from the Chernobyl exposure, it is well documented that a dramatic increase in childhood thyroid cancer has occurred in regions where significant exposure occurred. One recent report showed that since the Chernobyl accident, the incidence of thyroid cancer in 9-year-olds increased 50-fold in the "high exposure area" compared to an increase of 6-fold among 17-year-old children (69). Thus, although none of these results are specific to individual years of age, collectively they indicate a pattern of higher risk for radiation-induced thyroid disease at younger ages relative to adult ages.

Although there are few human studies of exposure to ^{131}I which can adequately evaluate the effect of age at exposure, animal studies have suggested greater carcinogenic risk at younger ages. In ^{131}I uptake experiments in rats, Sikov (98) demonstrated that fetal thyroids were 20 times more sensitive to functional damage than adult thyroids. Corresponding estimates for neonates and weanlings were 3 times and 1.5 times the sensitivity of adult thyroids, respectively. Similar results have been observed in guinea pigs (99). Christov (100) has reported similar findings using external radiation (x-rays) in Wistar rats. Among those irradiated at ten days of age, 40% developed thyroid adenomas whereas only 15% developed these tumors when irradiated at 60 days of age. None of the control rats developed tumors.

Although there are no human data to support age at exposure as an important factor in hyperparathyroidism after ^{131}I exposure, animal studies do provide some evidence that there is an increased frequency of parathyroid tumors in rats exposed to ^{131}I at young ages relative to older ages (80,81).

Based on these data from animal and epidemiological studies, it seems reasonable to expect that the risk of radiation-induced thyroid disease (and possibly hyperparathyroidism) would be greater among those exposed at the youngest ages. Therefore, the Pilot Study was limited to people who were children (ages 0-5) during the periods of greatest atmospheric releases from Hanford (i.e., 1945-1946). Thus, people born from 1942-1946 were eligible for inclusion in the Pilot Study.

This approach was also advantageous regarding important aspects of the fieldwork. Since the primary information to be used in the dosimetry calculations was to be derived from interviews with mothers (or other close relatives or individuals knowledgeable of the participant's childhood), it was important to maximize the probability that such information could be successfully collected and that it would be reasonably accurate. Younger study participants would, in general, have younger parents. Given an average follow-up of about 40 years, the 0-5 year age range at the time of exposure would reasonably assure that most parents of study participants would still be living and able to participate in an interview.

Following the Pilot Study, in order to include greater numbers of participants with thyroid doses in the higher range, it was also necessary to change the years of birth from which potential participants were selected (see section V.A.3 below). Given that the largest exposures would have occurred in 1945, it was thought more advantageous to select births from earlier years, 1940 and 1941, than from years later than those already included in the study. Thus, births from 1940 through 1944 were included in the Full Study selections. While there was some concern that including earlier births would decrease the numbers of participants for whom a CATI respondent could be found, this was not felt to outweigh the need to include as many higher dose participants as possible.

A.3. Other Possible Criteria

It is well established that thyroid neoplasia occurs more frequently in women (30), and there is evidence to suggest an increased risk among the Jewish (30,101). However, no attempt to further restrict eligibility in the Pilot Study based upon sex or ethnicity was made.

Although there are no Reservations in Benton, Franklin, or Walla Walla counties, Native American populations from the region were to be considered in the Pilot Study in an attempt to better define the radiation doses these populations may have received. Most Native Americans did not live in the areas around the Hanford Site where the highest thyroid doses were likely to have occurred. However dietary and/or lifestyle practices specific to one or more of the Tribes and Nations in the region may have been important in contributing to a radiation dose to the thyroid from Hanford's ^{131}I. Thus, as part of the Pilot Study, it was planned to attempt to determine whether the Native American populations in the region experienced exposures to radioactive iodine that could have resulted in significant thyroid doses. Section VII.A. specifies in more detail the conduct of this portion of the study. No attempt was made, however, to exclude people of Native American heritage from participation in the study.

B. Definition of Evaluable Participant

An evaluable participant was defined as one who could be located, who agreed to participate in the study, and for whom sufficient information could be obtained concerning both radiation exposure and thyroid outcomes. For each located participant, every attempt was made to obtain information from all possible sources regarding radiation exposure from Hanford. However, all information of possible use to this study was often not available, especially in view of the length of time that had elapsed since the years of peak exposure. Therefore, for living participants, sufficient information was defined as the following: In-Person Interview and physical examination. Although participants were asked to provide a blood sample and to receive an ultrasound examination (see section V.F, below), those who refused either or both were not deemed non-evaluable. Each participant's final assessment of thyroid or parathyroid disease status was based on the best historical and current information available (as described more fully in sections V.H-V.I).

For deceased subjects, sufficient information was defined as a residence history collected through a surrogate respondent, medical history from a surrogate, and medical record confirmation of thyroid disease reported by a surrogate. It was planned that persons who could serve as surrogate respondents for deceased subjects could include (but were not limited to): a parent, sibling, aunt or uncle of the subject. During the course of the study, it was determined that the plan for conducting CATIs with surrogates for the deceased subjects was not feasible. A discussion of the results of field tests of this portion of the study is contained in section V.D.4 below.

C. Outcome Criteria

This section describes the diagnostic criteria for the thyroid and parathyroid outcomes that are used in this study. Each outcome had two components: 1) the criteria established by the HTDS Study Management Team for the diagnosis of each outcome and 2) an indication of the basis for each diagnosis which serves as a measure of the quality of that diagnosis. The final diagnosis for each outcome included both the presence or absence of the diagnosis and if present, information about the basis of the diagnostic information**.** For example, information about the basis of the diagnosis included whether the diagnosis was made from the HTDS clinic evaluation, obtained from prior medical records with supporting documentation, obtained from prior medical records without supporting documentation, or obtained from a report by the participant or his or her Computer Assisted Telephone Interview (CATI) respondent, without documentation from either the HTDS evaluation or any prior medical records. Diagnostic information obtained from the HTDS evaluation and diagnostic information which was well documented in medical records and met criteria for HTDS diagnoses was considered to be the most definitive and of the highest quality. The primary analysis for each disease outcome was therefore restricted to cases defined according to these two sources. However, additional analyses were performed for each disease outcome using alternative definitions that were more inclusive and less definitive. These alternative definitions are provided in sections IX.C through IX.O below. If a participant had multiple sources of diagnostic information for a particular thyroid disease, with more than one basis for diagnosis, then he or she was classified according to the basis providing the most definitive diagnosis.

C.1. Thyroid Cancer

Diagnostic criteria: Thyroid malignancy according to histopathology reports from a surgical specimen. Original pathology slides were reviewed by the HTDS pathologist, whether the diagnosis was made by HTDS physicians or whether the participant had already had prior thyroid surgery.

Basis for diagnosis. The following categories indicate the basis for the diagnosis of thyroid cancer:

1. Diagnosis originating from HTDS evaluation based on subsequent histology
2. Diagnosis from prior medical record with documentation of histology
3. Clinical diagnosis from HTDS evaluation (no histology available)
4. Clinical diagnosis from prior medical record (no histology available)
5. Participant/respondent report only

C.2. Benign Thyroid Nodule

Diagnostic criteria: Any confirmed documentation of benign histology or cytology as interpreted by a pathologist.

Basis for diagnosis. The following categories indicate the basis for the diagnosis of benign thyroid nodule:

1. Histologic or cytologic diagnosis based on HTDS clinic evaluation
2. Histologic or cytologic diagnosis based on prior medical record documentation
3. Clinical diagnosis from either HTDS evaluation or medical records (clinical impression without cytology)
4. Participant/respondent report only

C.3. Any Thyroid Nodule

Diagnostic criteria: Any thyroid nodule which has been classified as thyroid cancer, a benign thyroid nodule, or a nodule which is suspicious for malignancy or neoplasm. The latter category represents nodules that have cytology suspicious for either malignancy or follicular neoplasm for which no surgery was performed and therefore no histology was available.

Basis for diagnosis. The following categories indicate the basis for the diagnosis of any thyroid nodule:

1. Histologic or cytologic diagnosis based on HTDS clinic evaluation
2. Histologic or cytologic diagnosis based on prior medical record documentation
3. Clinical diagnosis from either HTDS evaluation or medical records (clinical impression without cytology)
4. Participant/respondent report only

C.4. Hypothyroidism

Diagnostic criteria: Elevation of TSH above the upper limit of normal (5.0 μIu/ml) with either low or normal thyroid hormone levels.

Basis for diagnosis. The following categories indicate the basis for the diagnosis of hypothyroidism:

1. HTDS laboratory evaluation
2. Medical records with supporting documentation (elevated TSH)
3. Medical records without supporting documentation
4. Inferred from past or current thyroid hormone therapy
5. Participant/respondent report only

C.5. Autoimmune (Hashimoto's) Thyroiditis

Diagnostic criteria: Positive antithyroid antibody result on either antimicrosomal antibody (AMA) or anti-thyroperoxidase antibody (anti-TPO). Levels above the normal limits for these antibodies (AMA, greater or equal to 25 u/ml; anti-TPO, greater or equal to 2.0 Iu/ml) were considered positive. Participants with positive antibodies but with documentation of Graves disease were not included in this outcome category. Anti-thyroglobulin antibody was also used as an additional antibody marker for an alternative diagnosis of autoimmune thyroiditis (positive result: greater or equal to 1.0 Iu/ml).

Basis for diagnosis. The following categories indicate the basis for the diagnosis of autoimmune thyroiditis:

1. HTDS laboratory evaluation
2. Medical records with supporting documentation (positive anti-thyroid antibodies)
3. Medical records without supporting documentation
4. Participant/respondent report only

C.6. Graves Disease

Diagnostic criteria: Hyperthyroidism present (see # 8 below) with the following additional criteria:
1. Elevated radioiodine uptake and/or thyroid nuclear scan consistent with Graves disease; and/or
2. Exophthalmos

Basis for diagnosis. The following categories indicate the basis for the diagnosis of Graves disease:
1. HTDS laboratory and nuclear medicine evaluation
2. Medical records with supporting documentation
3. Medical records without supporting documentation
4. Participant/respondent report only

C.7. Autoimmune Thyroid Disease

Diagnostic criteria: Defined as having the diagnosis of either autoimmune thyroiditis or Graves disease. See above for diagnostic criteria and basis of diagnostic information for each of these outcomes.

8. **Basis for diagnosis.** In general, a diagnosis of autoimmune thyroid disease was simply assigned on the basis for diagnosis of the autoimmune thyroiditis or Graves disease that the participant had. In a small number of instances, participants had diagnoses of both autoimmune thyroiditis based on the HTDS laboratory evaluation or medical records with supporting documentation, and of Graves disease based on medical records without supporting documentation or on participant/respondent report only. In all of these instances, the basis for the diagnosis of autoimmune thyroid disease was taken to be the more definitive, i.e. HTDS laboratory evaluation or medical records with supporting documentation.

C.8. Hyperthyroidism

Diagnostic criteria: Suppressed TSH (less than 0.32 μIu/ml) in the presence of normal or high thyroid hormone levels. The following additional information was collected to further assess the etiology of hyperthyroidism:
1. To evaluate Graves disease or a toxic thyroid nodule as an etiology of a suppressed TSH, repeat thyroid function tests (TSH, T3 and T4 levels), a thyroid nuclear scan and radioiodine uptake were requested (see #6 above).
2. History of current medical treatment with thyroid hormone was obtained to assess exogenous thyroid hormone therapy as a cause of hyperthyroidism. For participants having a suppressed TSH while taking thyroid hormone medication, their hyperthyroidism was presumed to be caused by exogenous thyroid hormone.

Basis for diagnosis. The following categories indicate the basis for the diagnosis of hyperthyroidism:
1. HTDS laboratory evaluation
2. Medical records with supporting documentation
3. Medical records without supporting documentation
4. Participant/respondent report only

C.9. Multinodular Thyroid Gland

Diagnostic criteria: A thyroid gland with abnormal firm consistency with two or more discrete nodules, or multiple firm lobular and/or nodular areas throughout the gland. The definition of thyromegaly in this study is a two-fold enlargement of the thyroid gland based on physical examination. Therefore, the above characteristics of a multinodular gland occurring in a gland enlarged two-fold or more is classified as multinodular goiter whereas these characteristics occurring in a gland of normal size (less than two-fold enlarged) is classified as multinodular gland. The definition of thyromegaly as a two-fold increase in thyroid gland size was chosen as a conservative definition to avoid classifying normal variations as clinical disease. Dominant palpable nodules or those which were nonpalpaple and greater than 1.5 cm in three dimensions underwent FNA biopsy. Such nodules would then be classified as either benign or malignant depending on the results of the biopsy or further thyroid surgery.

Basis for diagnosis. The following categories indicate the basis for the diagnosis of multinodular gland:

1. HTDS physical examination
2. Medical records with documentation of multinodular gland or goiter
3. Participant/respondent report only

C.10. Simple Goiter

Diagnostic criteria: Diffuse thyromegaly (two-fold enlargement) with normal consistency and without palpable nodules or lobulations. The definition of thyromegaly as a two-fold increase in thyroid gland size was chosen as a conservative definition to avoid classifying normal variations as clinical disease. This classification was intended primarily to reflect physiologic thyroid gland enlargement.

Basis for diagnosis. The following categories indicate the basis for the diagnosis of simple goiter:

1. HTDS physical examination
2. Medical records with documentation of diffuse goiter without nodularity or abnormalities in consistency
3. Participant/respondent report only

C.11. Other Thyroid Disease

Diagnostic criteria: This category was designated for any diagnoses of thyroid disease that are not included in the HTDS diagnostic outcomes above. It was primarily a category for participant reports of unknown thyroid disease, diagnosed generally many years ago, and treated with unknown therapy for which no medical records were available.

Basis for diagnosis. The following categories indicate the basis for the diagnosis of other thyroid disease:

1. HTDS evaluation
2. Participant/respondent report only

C.12. Ultrasound-Detected Abnormalities of the Thyroid (Thyroid UDAs)

Diagnostic criteria: The following categories of ultrasound abnormalities were defined:

1. Palpable ultrasound-detected thyroid abnormalities
2. Nonpalpable focal ultrasound-detected thyroid abnormalities
3. Diffuse (nonpalpable) ultrasound-detected abnormalities of the thyroid
4. Any ultrasound-detected abnormality of the thyroid (any of the above categories)

Basis for diagnosis. All of these definitions were based on only one source of information: HTDS ultrasound examination.

C.13. Hyperparathyroidism

Diagnostic criteria: Defined as hypercalcemia (calcium greater than 10.2 mg/dl) with an elevated PTH level (greater than 65pg/ml).

Basis for diagnosis. The following categories indicate the basis for the diagnosis of hyperparathyroidism:

1. HTDS laboratory evaluation
2. Medical records with supporting documentation
3. Medical records without supporting documentation
4. Participant/respondent report only

V. FIELD PROCEDURES AND METHODS, RESULTS OF DATA COLLECTION PROCESS

A. Cohort Definition, Subject Identification and Selection

A.1. Background

A.1.a. Objectives

The objective of this component of the study was to define and identify a group of people (a cohort) who were exposed to atmospheric releases of radioactive iodine (^{131}I) from the Hanford Nuclear Site between 1944 and 1957. Since the primary objective of the overall study was to determine whether exposure to such radiation resulted in an increased risk of thyroid disease, it was important to identify a cohort within which there would be the greatest likelihood of detecting an association between exposure to ^{131}I from Hanford and thyroid disease, if such a relationship exists. This was to be accomplished by defining a cohort that would contain adequate numbers of people with the highest possible radiation doses to the thyroid from Hanford, as well as people with very low radiation doses to the thyroid from Hanford.

A.1.b. Definition of the Cohort

In seeking to define a cohort that would contain individuals with a full range of exposures to ^{131}I from Hanford, extensive attempts were made to investigate different sources of information that would enable one to construct a comprehensive list of people who might have been exposed. Ideally, such a list would consist of all people in a relatively large population surrounding the Hanford site who were resident during the time period that the largest atmospheric releases occurred, and would contain enough identifying information to ensure that a sufficient number of people could be located nearly five decades after exposure. The following sources of information were investigated in the Hanford region: 1) school enrollment records; 2) school health records; 3) school reunion lists; 4) health department clinic and immunization records; 5) church membership lists; 6) town lists and voter registration records; 7) Census Bureau records; 8) Internal Revenue Service records; 9) property tax and public utility records; and 10) birth records.

Most of these sources of information proved to be inadequate for constructing a sufficiently comprehensive listing of individuals who might have been exposed to Hanford releases. School health records, reunion lists, health department records, church lists, town and voter lists, and property and utility records were all too incomplete. School enrollment records were complete and potentially very useful where they existed, but unfortunately many school districts in the region had destroyed old records and a few denied us access. Census Bureau and IRS records would have been ideal sources for enumerating a population, but access to such information was prohibited by law. Thus, only birth records provided an acceptable source for identifying a cohort.

Birth records provide a complete listing of all people born in a defined geographic area during defined time periods. The records were available at no cost to the study and could be easily accessed by staff. Thus, by abstracting information directly from birth certificates, it was possible to construct a roster of individuals corresponding to specific geographic areas and time periods most relevant to the Hanford releases.

A.2. Plan

A.2.a. Protocol Plan

For the Pilot Study, a birth roster was constructed based on all birth certificates from the counties of Benton, Franklin, Walla Walla, Okanogan, Ferry, and Stevens for the years 1942-1946. As indicated above, complete birth records existed for these counties and were available from the State of Washington Vital Records Division. The following data were abstracted from each birth certificate and entered into a computerized database to form the roster for selection of potential participants: birth certificate number, mother's usual residence, child's name, sex, and birthdate, father's name, mother's name, mother's mailing address, and county of birth.

The field "Mother's Mailing Address" was judged to best indicate the mother's actual residence when the subject was born. However, the birth certificates for births in Benton, Franklin, and Walla Walla counties had been computerized previously for the CDC by the State of Washington, and "Mother's Mailing Address" was not included in this database. Thus only "Mother's Usual Residence" was available and it was felt this might not reflect the mother's actual residence when the subject was born. Therefore, HTDS staff computerized the mailing addresses for those mothers who gave birth in Benton, Franklin, and Walla Walla Counties but whose usual residences were outside the six study counties in order to include those whose "Mother's Mailing Address" lay within the six Pilot Study counties in the roster for subject selection.

For purposes of geographical stratification, "Mother's Residence at the Subject's Birth" was defined for the counties of Benton, Franklin and Walla Walla as follows:

- For births with "Mother's Usual Residence" (birth certificate item 2) in one of the six Pilot Study counties (Benton, Franklin, Walla Walla, Okanogan, Ferry and Stevens), "Mother's Residence at Subject's Birth" was defined as the "Mother's Usual Residence"

- For births with "Mother's Usual Residence" outside the six study counties, "Mother's Residence at the Subject's Birth" was defined to be the "Mother's Mailing Address".

For Okanogan, Ferry and Stevens counties, "Mother's Residence at the Subject's Birth" was defined as the "Mother's Mailing Address."

In addition, birth records for Spokane and Yakima counties were reviewed to ascertain births that occurred in those counties to residents of Benton, Franklin, and Walla Walla counties. "Mother's Residence at the Subject's Birth" for these certificates was also assigned to be the "Mother's Mailing Address." People for whom the "Mother's Residence at the Subject's Birth" was outside the selected counties were excluded from the roster.

Eligibility for the study was limited to people whose "Mother's Residence at the Subject's Birth" was in one of the selected counties.

A.2.a.1. Rationale

As noted in section IV-A above**,** preliminary findings from the HEDR project regarding meteorological conditions affecting the deposition and concentration of radioactive iodine in vegetation, and the patterns of milk production and consumption by county, indicated that people with the highest thyroid doses were most likely to have lived in the area encompassed by Benton, Franklin, and Walla Walla counties. Thus, in the Pilot Study for the purposes of subject selection only, residence at time of birth acted as a surrogate for the anticipated radiation dose to the thyroid from Hanford. Individual thyroid radiation dose could only be estimated from data collected during the study. The selection of cohort members was also extended to include three counties on the Canadian border north of the Hanford site

(Okanogan, Ferry and Stevens). These counties were selected because, based upon the information available at the time regarding possible radiation doses to the thyroid, they could be expected to contribute some cohort members with very low radiation doses to the thyroid from Hanford. In addition, people living in these counties would likely be comparable to those who receive higher thyroid doses in terms of other factors which could potentially influence the risk of thyroid disease (e.g., geography, urban/rural composition, occupational factors, socioeconomic factors, age, ethnicity, sex). Furthermore, similar opportunities and resources existed to identify and trace people in this group as in the group that received a thyroid dose. Thus, a cohort was selected which was expected to contain people whose dose estimates would range from the highest doses received to the lowest.

Preliminary estimates of the HEDR project suggested that the highest thyroid doses were probably in people exposed as infants or children during the first years of Hanford operations. This is because infants and children receive higher thyroid doses per unit exposure due primarily to the small size of their thyroid glands. In addition, existing literature suggests that the risk of radiation-induced thyroid disease (and possibly hyperparathyroidism) is greatest among those exposed at youngest ages (see section II.B. for a more detailed description). For these reasons, the Pilot Study was limited to people born from 1942-46, since the large majority of releases of radioactive iodine from the Hanford facility occurred in 1944-46 (with the exceptions of the "Green Run" in December 1949 and the releases during 1951). Thus, the cohort would contain people whose exposures began as early as the prenatal period, and as late as age three. An additional benefit of choosing this group was that mothers and close relatives of people born during 1942-46 would more likely be alive and available for interview compared to those of people born earlier.

Selection of potential participants from the Birth Roster was stratified by geographical area, year of birth, and sex. The purpose of stratification by geographical area and birth year was to assure that adequate numbers of high dose and low dose participants were identified, and a wide range of doses was obtained. Stratification by sex also reduced the possibility of confounding by sex that could reduce the efficiency of the study.

For purposes of stratified selection of subjects from the Birth Roster, geographical areas were defined to distinguish predominantly rural areas from predominantly urban areas. The reason for such distinction was that it was reasonable to expect that people from predominantly rural areas may have been more likely to consume fresh raw milk than their more urbanized counterparts. If true, and if such consumption patterns were an important determinant of higher dose, it might be important in the Full Study to concentrate potential participant selection from rural areas. At the time of protocol development, HEDR Phase I results indicated that the distinction between fresh raw and commercial milk consumption did not have a substantial effect on the magnitude of estimated thyroid dose (125).

Each person on the Birth Roster was assigned to the area which contained his or her "Mother's Residence at the Subject's Birth," as outlined in section V.A.2.a. Eight geographical areas, called "geostrata" in this report, were defined:

1. Richland
2. Pasco/Kennewick
3. Walla Walla City
4. Benton County outside Richland and Kennewick
5. Franklin County outside Pasco
6. Walla Walla County outside Walla Walla City
7. Okanogan County
8. Ferry and Stevens Counties

A.2.a.2. Completeness Required for Success

For each of the eight geostrata defined above, a target of ten living evaluable participants was sought for the Pilot Study for each sex and year of birth. For geostrata 2-6, there were ten strata (five years of birth x two sexes) for a total target of 100 living evaluable participants in each area. For geostratum 1 (Richland), there were six strata (three years x two sexes, as Richland was not defined as a geostratum prior to 1944) for a total target of 60 living evaluable participants. A target of five living evaluable participants was sought in each of the ten year/sex strata for geostrata 7 and 8 (for a total of 100), however during the Pilot Study sample selection the target of 10 living evaluable participants was actually used, for a total of 200). Thus, the Pilot Study attempted to enroll no less than 560 participants from geostrata 1-6, and 200 participants from geostrata 7-8. As a first approximation, twice this number of subjects was to be selected from the Birth Roster to obtain the overall goal of 760 living evaluable participants. The plan was that if this goal was not achieved (i.e., less than a 50% success rate in locating and enrolling participants), additional subjects would be selected in the same manner.

A.2.b. Plans for Assessing the Need for Change in the Full Study

The feasibility of basing the Full Study on a cohort identified solely from birth certificates depended in part on whether adequate information could be obtained for a sufficiently large proportion of cohort members, and whether the range of thyroid radiation doses obtained was sufficiently wide. However, any decision regarding the roster of subjects for the Full Study (e.g., whether to include additional birth year cohorts or participants identified from other sources, such as school records) would be based on all pertinent information, and not just the data obtained and used to evaluate the above criteria.

It was anticipated that the birth cohort criteria for defining cohort members in the areas most heavily exposed would be expanded for the Full Study. Thus, assuming the same methods for identifying cohort members (i.e., birth certificates), it was expected that additional birth year cohorts might be included. Such an expansion would likely be achieved by including people born before 1942, as this would continue to provide the best opportunity to include people who received relatively high doses during childhood. It would also serve to include a larger range of ages at exposure. However, it might also be possible to expand the range of birth years slightly forward in time as well.

Decisions about whether and how to expand the cohort were to be based largely upon the sample size calculations conducted at the conclusion of the Pilot Study and the resources available to the study. If insufficient numbers of births were available under the current criteria to satisfy the sample size requirements of the Full Study, then clearly it would almost certainly be necessary to expand the cohort to include additional births from other years. If, however, it was not necessary to expand the cohort to meet sample size requirements, such expansion, to the extent that resources allowed, would nevertheless be proposed to increase the generalizability of the results by including a wider population representation in the study. Secondarily, such an expansion would serve to increase the power of the study.

At the time the protocol was written, it was unclear whether the geographical boundaries of the study area would change. It was considered unlikely that the boundaries of the area exposed would be significantly expanded. More likely, it was thought that it might be possible (and advantageous) to restrict the definition of "exposed" areas somewhat, based on the distribution of preliminary doses observed in the Pilot Study.

To determine whether a geographically separate area should be identified, it was planned to evaluate: 1) the dose distributions for participants born in the northern three counties (Okanogan, Ferry, and Stevens); and 2) the degree to which all aspects of data collection among people geographically removed from the Hanford site (and presumed to be less likely to be highly exposed) relative to those in closer proximity was successfully conducted. As described in section III-J.1 of the protocol, collectively, these evaluations would allow a better determination of whether the geographically separate areas chosen

for the Pilot Study would be suitable regarding doses (i.e., that most participants in those geostrata would have relatively low doses) and logistics. It could be, for example, that it would not be necessary to include all of these counties in the Full Study. In contrast, it was also recognized that the results of the Pilot Study could indicate that none of the separate counties would be suitable for use in a Full Study. If such a determination was made based upon dose distributions (and not issues of feasibility of data collection), it would be necessary to explore and define another geographical region or regions more removed from the Hanford Site to maintain the capability of being able to conduct analyses that were not solely dependent on HEDR individual dose estimates. The evaluation of other potential regions would be based primarily on the following factors: 1) meteorological data, 2) milk distribution patterns, and 3) socioeconomic and lifestyle factors.

A.3. Revisions

A.3.a. Rationale for Revisions made in the Transition Sample

Prior to completion of the Pilot Study and before a final determination had been made regarding the conduct of a Full Study, another selection of cohort members was made from the Birth Roster. This was done after consultation with the CDC and the HTDS Federal Advisory Committee, in anticipation of continuing with a Full Study, to maintain continuity in field operations and study personnel. This group, called the Transition Sample, was selected prior to any analyses of the dose data from the Pilot Study. The Transition Sample was selected from each birth year and sex stratum in each of the following geostrata: Richland, Pasco/Kennewick, Walla Walla City, Benton County, and Walla Walla County (further selection from the Franklin County geostratum was not possible since all subjects in that geostratum had already been selected for the Pilot Study sample). The Transition Sample was selected from these geostrata because they were the most likely to have relatively high doses, and it was felt there were already sufficient numbers of low dose participants.

A.3.b. Rationale for Revisions made in the Full Study

The power calculations described in detail in the Pilot Study Report outline the rationale for the revisions made in the Full Study selection (see section V.A-5 below for a brief summary of these calculations). In short, it was determined that the cohort defined for the Pilot Study was likely to be inadequate in size, and that greater numbers of participants in the higher dose range would be needed to ensure sufficient statistical power for the primary dose-response analyses. Therefore, in order to include greater numbers of participants with thyroid doses in the higher range, it was necessary to change the years of birth from which cohort members were selected. Given that the largest exposures would have occurred in 1945, it was thought more advantageous to select births from earlier years, 1940 and 1941, than from years later than those already included in the study. Thus, births from 1940 through 1944 were included in the Full Study selections. While there was some concern that including earlier births would decrease the numbers of participants for whom a CATI respondent could be found, this was not felt to outweigh the need to include as many higher dose participants as possible.

Essentially final results of the HEDR project became available while the HTDS Pilot Study was in progress. These results suggested that the geographical region defining the HTDS cohort should be revised to meet the objective of including as many people with the highest thyroid doses as possible. In particular, the final HEDR results suggested that people born in Adams County might be more likely to have higher thyroid doses from Hanford than those in the Walla Walla geostrata.

The Full Study cohort was therefore defined initially to include the Pilot Study and Transition Samples along with (1) all remaining 1942-44 births in the Richland, Pasco/Kennewick, Benton County and Franklin County geostrata; (2) all 1940-41 births with mother's residence at subject's birth in Benton or Franklin County (which include the Pasco/Kennewick strata); and (3) all 1940-44 births with mother's

residence at subject's birth in Adams County. These birth years and geostrata were selected to ensure the inclusion of more high dose participants, based on the dose estimates for hypothetical representative individuals in the HEDR final report of April 21, 1994.

The definition of the cohort was expanded one time, after it was determined that the number of birth certificates obtained was lower than had been projected. Originally it had been projected that 3427 living evaluable participants would be found in the Full Study cohort. However, after obtaining the birth certificates, because there were fewer births than anticipated and the Pilot Study showed there would be fewer living evaluable participants than originally estimated, this projection was reduced to 3006. Therefore, the cohort was expanded to include all remaining births between 12/31/44 and 6/30/45 in the Richland, Pasco/Kennewick, Benton, and Adams County geostrata (no births in the Franklin County geostratum remained unselected). This increased the projected number of living evaluable participants from 3006 to 3277, thereby maintaining essentially the same levels of power as originally projected.

A.4. Outcome and Final Results

The final definition of the cohort was as follows:

1. All births from 01/01/40 to 06/30/45 (in the study counties searched for occurrence births) with mother's residence at subject's birth in Benton, Franklin, or Adams counties (including the Richland and Pasco/Kennewick geostrata).
2. A randomly selected subset of births from 07/01/45 to 12/31/46 in the same counties with mother's residence at subject's birth in Benton or Franklin counties (including the Richland and Pasco/Kennewick geostrata).
3. A randomly selected subset of births from 01/01/42 to 12/31/46 with mother's residence at subject's birth in Walla Walla County (including the Walla Walla geostratum) or Okanogan, Ferry, or Stevens counties.

Table V.A-1, below, shows the birth years within each geostratum from which subjects were selected, by each phase of selection. Three separate selections were conducted to complete the Full Study Sample, after the Pilot Study and Transition Samples were selected. Table V.A-2 shows the numbers of participants selected in each of the 100 strata for the Full Study. Note that all people in the 1940-1944 birth cohorts for Benton, Franklin and Adams Counties (including the Richland and Pasco/Kennewick geostrata) were selected.

Table V.A-1. Birth Years Included in Each Phase of Participant Selection

	Phase of Selection				
Geographic Area	Pilot	Transition	Full 1	Full 2	Full 3
Richland	1944-46*	1944-46	1944		1/45-6/45
Pasco/Kennewick	1942-46	1942-46	1942-44	1940-41	1/45-6/45
Walla Walla City	1942-46	1942-46			
Benton County	1942-46*	1942-46	1942-44	1940-41	1/45-6/45
Franklin County	1942-46			1940-41	1/45-6/45
Walla Walla County	1942-46	1942-46			
Okanogan County	1942-46				
Ferry/Stevens Counties	1942-46				
Adams County				1940-44	1/45-6/45

* The city of Richland was defined as a geostratum separate from Benton County beginning in 1944.

Table V.A-2 Distribution of Birth Year, Sex, and Geostratum for the Full Study Cohort

		Birth Year														
		1940		1941		1942		1943		1944		1945		1946		
Geostratum		F	M	F	M	F	M	F	M	F	M	F	M	F	M	Total
Richland*	Births									92	93	234	230	237	197	1083
	Selected									92	93	142	128	43	44	542
	% Selected									100	100	60.7	55.7	18.1	22.3	50.0
Pasco/	Births	63	77	84	82	84	83	140	162	216	228	209	209	243	204	2084
Kennewick	Selected	63	77	84	82	84	83	140	162	216	228	131	127	41	40	1558
	% Selected	100	100	100	100	100	100	100	100	100	100	62.7	60.8	16.9	19.6	74.8
Walla Walla	Births					179	205	184	182	260	255	300	336	307	322	2530
(city)	Selected					44	41	41	40	40	42	40	42	41	41	412
	% Selected					24.6	20.0	22.3	22.0	15.4	16.5	13.3	12.5	13.4	12.7	16.3
Benton	Births	75	47	69	71	52	71	85	86	176	187	60	67	72	75	1193
County*	Selected	75	47	69	71	52	71	85	86	176	187	51	57	48	50	1125
	% Selected	100	100	100	100	100	100	100	100	100	100	85.0	85.1	66.7	66.7	94.3
Franklin	Births	19	19	7	20	22	17	12	13	23	20	11	14	15	22	234
County	Selected	19	19	7	20	22	17	12	13	23	20	11	14	15	22	234
	% Selected	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Walla Walla	Births					47	53	46	66	66	71	71	58	84	80	642
County	Selected					47	53	46	44	44	48	48	54	42	40	466
	% Selected					100	100	100	66.7	66.7	67.6	67.6	93.1	50.0	50.0	72.6
Okanogan	Births					217	236	255	280	222	237	227	253	317	306	2550
County	Selected					21	22	21	20	21	22	21	21	21	21	211
	% Selected					9.7	9.3	8.2	7.1	9.5	9.3	9.3	8.3	6.6	6.9	8.3
Ferry/	Births					231	233	215	234	178	197	125	117	198	227	1955
Stevens	Selected					21	21	22	21	22	22	21	24	22	20	216
Counties	% Selected					9.1	9.0	10.2	9.0	12.4	11.2	16.8	20.5	11.1	8.8	11.0
Adams**	Births	30	31	37	36	37	44	45	44	48	45	17	21			435
County	Selected	30	31	37	36	37	44	45	44	48	45	17	21			435
	% Selected	100	100	100	100	100	100	100	100	100	100	100	100			100
Total	Births	187	174	197	209	869	942	982	1067	1281	1333	1254	1305	1473	1433	12706
	Selected	187	174	197	209	328	352	412	430	682	707	482	488	273	278	5199

* The city of Richland was defined as a geostratum separate from Benton County beginning in 1944.
** 1945 Adams County number of births is for January-June only (all other geostrata include some July-December 1945 births).

A.5. Summary of Full Study Power Calculations, as Presented in HTDS Pilot Study Report

Two primary objectives of the Pilot Study were to assess the suitability of areas chosen for the selection of study participants, and to utilize Pilot Study dose information and response rates to estimate sample sizes required to achieve adequate statistical power for a Full Study. As is often the case with observational studies such as the HTDS, sample size and dose distribution cannot be chosen independently. In particular, as the sample size increases, the relatively small groups of subjects likely to have the highest doses are all selected, and further selections must be made from the relatively larger groups of people likely to have smaller doses. As a result, beyond a certain number, the effect of increasing sample size is to a certain extent offset by the effect of decreasing mean and variance of the resulting dose distribution.

As described in Appendix H of the HTDS Protocol (1), the primary power calculations focused on tests of the dose-response for the endpoint of thyroid neoplasia (malignant and benign). Calculations were also performed for two additional endpoints: thyroid malignancy and ultrasound-detected abnormality of the thyroid (thyroid UDA). These three outcomes were selected since they provided a range of baseline outcome percentages: low (malignancy), intermediate (thyroid neoplasia), and high (UDA). Sample sizes were calculated for the χ^2 test for linear trend in the cumulative incidence of disease with stratification by sex. In particular the sample size N required for the one-sided test with critical level α to achieve statistical power $1-\beta$ to detect a dose-response coefficient B is given by the formula:

$$N = \left(z_{1-\alpha} - z_{\beta}\right)^2 \Big/ \left[B^2 \sigma^2 \sum_{i=1}^{I} \frac{\pi_i}{P_i^*\left(1 - P_i^*\right)} \right]$$

where

$z_p = \Phi^{-1}(p)$ denotes the 100p-th percentile of the standard normal distribution,
$i = 1,2$ indexes the I=2 sexes,
π_i = proportion of the N participants of sex denoted by i,
σ^2 = variance of the dose distribution, and
$P_i = P_i(\mu) = A_i + B\mu$ is the probability of disease for sex denoted by i and dose equal to the mean dose μ.

This formula indicates that, as is typically the case, the required sample size is largely determined by the variance σ^2: in particular the required sample size is roughly inversely proportional to σ^2. The effect of the mean dose is much more limited. For a given sample size N, the equation above can be solved for the power $1-\beta$. This approach was used under various assumptions about sample size, and power was displayed in figures as a function of the dose-response coefficient B. The resulting plot indicates in a comprehensive way the power of the planned analyses to detect radiation effects of various magnitudes.

The approach taken in the analysis of the Pilot Study results was to project the mean and variance of doses that might be obtained under various plans for selecting subjects to complete the sample for a Full Study. Consideration focused on three such plans:

Plan 1:

Remaining sampling would be restricted to birth records from 1942-44 for the Richland, Pasco/Kennewick, Benton County, and Franklin County regions. All remaining subjects from these strata would be included in the Full Study.

Plan 2:

In addition to Plan 1, the definition of eligibility would be expanded to include births during 1940-41 to mothers whose residence at time of birth was in Benton and Franklin Counties, and all such births would be included in the Full Study. Note that these two counties include Pasco and Kennewick, which do not need to be distinguished as a separate geographical region.

Plan 3:

In addition to Plan 2, the definition of eligibility would be expanded to include births during 1940-44 to mothers with residence at the time of birth in Adams County, and all such births would be included in the Full Study.

Note that Plan 1 required only projections based on dose data available from the Pilot Study, while Plans 2 and 3 required projections of dose distributions for years and/or regions not included in the Pilot Study. The methods for calculating projected means and variances for both types of projections were described in Appendix B of the Pilot Study Final Report.

The projected sample sizes, means and variances for these three plans, based on all dose data available from either the Pilot Study sample or the combined Pilot Study and transition sample, are shown in Table V.A-3.

Table V.A-3. Sample Size (N) and Projected Dose Mean and Variance (rad) of Full Study Dose Distribution for the Three Additional Sampling Plans

		Pilot Only (n=869)		Pilot and Transition (n=1139)	
Plan	N	Mean	Variance	Mean	Variance
1	2619	13.7	361.1	14.5	372.9
2	3081	13.5	353.6	14.4	367.3
3	3427	14.6	393.8	15.4	404.3

The HTDS was based on a cohort of people defined by the following eligibility criteria:

- Mother's residence at the time of the participant's birth: Benton, Franklin, Walla Walla, Okanogan, Ferry, Stevens, or Adams County in Washington State
- Year of birth: 1940 – 1946.

The rationale for this choice of counties and years is described in sections IV.A.1 and IV.A.2 below. The mother's usual residence at the time of the participant's birth, which can be determined from birth records, was used as a criterion since it was likely to indicate the participant's place of residence during the first years of Hanford's operations, when the largest releases of ^{131}I occurred (see section V.A.2 below). The cohort included the majority of the possible combinations of the seven counties and seven birth years. However birth year subcohorts for certain counties were not included since they were unlikely to include many participants with relatively high thyroid radiation doses (see sections V.A.2 and V.A.3 below).

Power functions of tests for dose-response based on the projections derived from the Pilot Study only are shown in Figures V.A-1 through V.A-3 for the endpoints of thyroid neoplasia (benign and malignant combined), thyroid cancer, and thyroid UDAs. In each figure, the lowest curve is based on Plan 1, and the highest on Plan 3.

Figure V.A-1. Projected Power Function: Thyroid Neoplasia Plans 1, 2 and 3

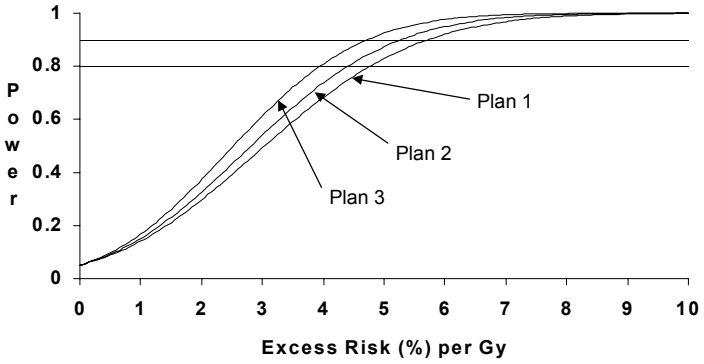


Figure V.A-2. Projected Power Function: Thyroid Malignancy Plans 1, 2, and 3

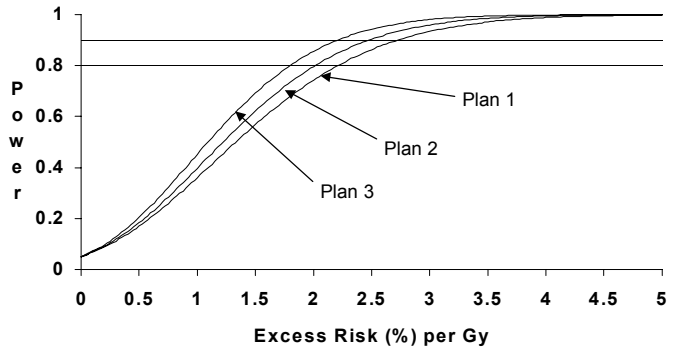
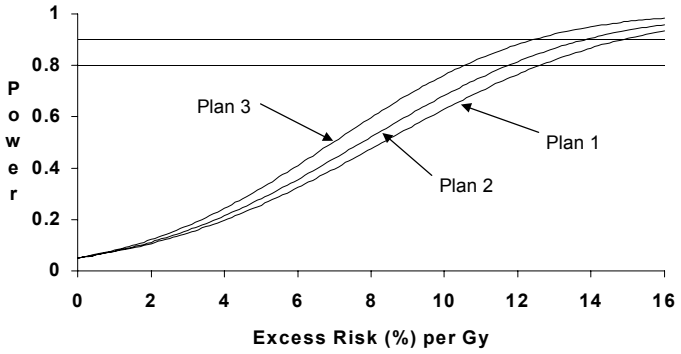


Figure V.A-3. Projected Power Function: Ultrasound Detected Abnormalities of the Thyroid Plans 1, 2 and 3



For the analysis of thyroid neoplasia (defined in the Protocol and Pilot Study as "all thyroid nodules)," the baseline percentages of patients with disease were taken as 5% for women and 2% for men (see Appendix H of the HTDS protocol for the derivation of these percentages). Based on data from the 869 Pilot Study participants, Plan 1 was projected to provide power of 0.83 to detect a dose-response coefficient of 5% per Gy. Under Plans 2 and 3 this increased to 0.87 and 0.92, respectively (Figure V.A-1). Thus with Plan 3 there would be adequate power to detect about a doubling (tripling) of risk among women (men) at 1 Gy (1000 mGy). This magnitude of effect is similar to that projected from the relative risk model of BEIR V as described in Appendix H of the HTDS protocol. It is also comparable to that recently reported for the Utah Study. Kerber et al. (53) reported a significant radiation dose-response for thyroid neoplasia during 1965-86 (p=0.019), with an estimated relative risk of 8.0 at 1 Gy (95% lower confidence bound 1.7).

The baseline percentages of participants with thyroid malignancy were taken to be 0.7% for women and 0.3% for men (see Appendix H of the protocol). Plans 1, 2, and 3 were projected to provide power of 0.80, 0.90, and 0.94, respectively, to detect a dose-response of 2.5% per Gy (Figure V.A-2).

For an analysis of thyroid UDAs, the baseline percentage of participants with such findings was taken as 40% for both sexes, based on information available from reports of thyroid ultrasound screening in unselected populations and the experience in the Pilot Study. Plans 1, 2, and 3 were projected to provide power of 0.90, 0.94, and 0.97 to detect a dose-response of 15% per Gy (Figure V.A-3).

A number of assumptions were made in the projections of statistical power described above. To assess the sensitivity of the projections to these assumptions, i.e., to assess whether deviations from any of these assumptions might lead to significant changes in the projected levels of power, additional power calculations were performed with these assumptions modified. The following assumptions were examined in these sensitivity calculations: 1) baseline rates of disease, 2) projected sample size, 3) doses from expanded In-Person Interviews, 4) doses for participants born during 1940-41, and 5) Adams County doses. In addition, the combined effects of deviations in more than one of these assumptions were investigated. The detailed results of these sensitivity calculations were provided in the Pilot Study Final Report (pages 55-74). Based on the results, the following conclusions were reached:

1. Cohorts identified from birth records were likely to provide a sufficiently wide distribution of doses for successful completion of a Full Study.

2. The cohorts defined for the Pilot Study were likely to be inadequate for completing a Full Study, and they should be augmented by the additions of 1940-41 Benton and Franklin Counties and 1940-44 Adams County births.

As described above, the modification proposed in the second conclusion was adopted. However, following the collection of the birth certificate data for the additional birth years and Adams County, and the analysis of more complete data regarding participation rates, it was apparent that further expansion of the cohort was needed. This was accomplished by extending the range of birth dates for Benton, Franklin, and Adams Counties to June 30, 1945.

B. Tracing Potential Participants

B.1. Background

The HTDS was conducted as a follow-up cohort study. Members of the study cohort were identified based on location of birth in the early to mid-1940's from birth certificates. Consequently, extensive effort would be required to locate cohort members, who were young children at the time of exposure, as adults nearly fifty years later. In addition, to identify all past and present thyroid disease in cohort members, participation in the study could not be limited to telephone contact, but would require in-person attendance for medical evaluation regardless of the participant's current area of residence.

B.1.a. Objectives of Tracing

The primary objective of the tracing was to identify a current address and telephone number for all living potential participants, so they could be recruited to participate in the study. A second objective was to obtain confirmation of death, as well as date and cause of death for all those deceased.

B.1.b. History of Tracing Efforts Around Hanford

Prior to the HTDS, a separately funded study had been conducted by investigators at the FHCRC to determine if former residents of the Hanford area could be traced to their current residences for the purposes of an epidemiologic study of radiation releases from the Hanford site. The primary objectives of this preliminary study were: 1) to design and test field procedures for identifying a group of potentially exposed persons; 2) to attempt to trace each person forward in time to the present or until death; 3) to obtain a current address and/or telephone number for each person; 4) to explore the feasibility of interviewing people identified; and 5) to explore the feasibility of obtaining medical records to verify self-reported illness histories.

The population selected for this preliminary study was defined by the rural area directly east of the Hanford Nuclear Site in Franklin County, containing 37 farm blocks subdivided into approximately 1897 farm units. Four farm blocks were randomly selected from the area, two being in the area of the two research interviewers' homes, and two being remote from these areas. In this manner, it was hoped that each interviewer would be working within an area that was very familiar and within which she would personally know the residents, as well as in one area which was quite unfamiliar.

Each interviewer was to obtain as much information as possible about anyone who resided in the assigned farm blocks from the time they were first inhabited until the present. Most farm blocks were first inhabited in the early 1950's due to the Columbia Basin Land Reclamation Project. However, some farm blocks were inhabited as early as 1909.

The principle sources of information the interviewers used to initially identify the population living in the selected farm blocks was a title company in Pasco, Washington which recorded a complete record of ownership for each farm unit. Because ownership records do not include family members or other residents, additional information was obtained through library references, telephone books, personal visits, and telephone calls. Thus, a chronology of persons who resided in each farm unit was constructed.

A total of 126 persons were found to have resided on the 14 occupied farm units within the four farm blocks. This number includes primary owners/residents, their children, employed farm workers and their families. For all but four residents (3%), actual years of residency were ascertained. Sixty-four (51%) residents lived on the farm units for some period between 1949 and 1965, a time period that encompassed much of the radioactive releases from Hanford.

Of the 126 persons identified, ten (8%) were documented to be deceased. An additional eight persons (6%) could not be located. Slightly more than half of those identified and located were currently still living on the farm units. Another 15% were resident in the immediate area, and 5% were within Washington or a neighboring state. The remainder of those located resided in the Western U.S.

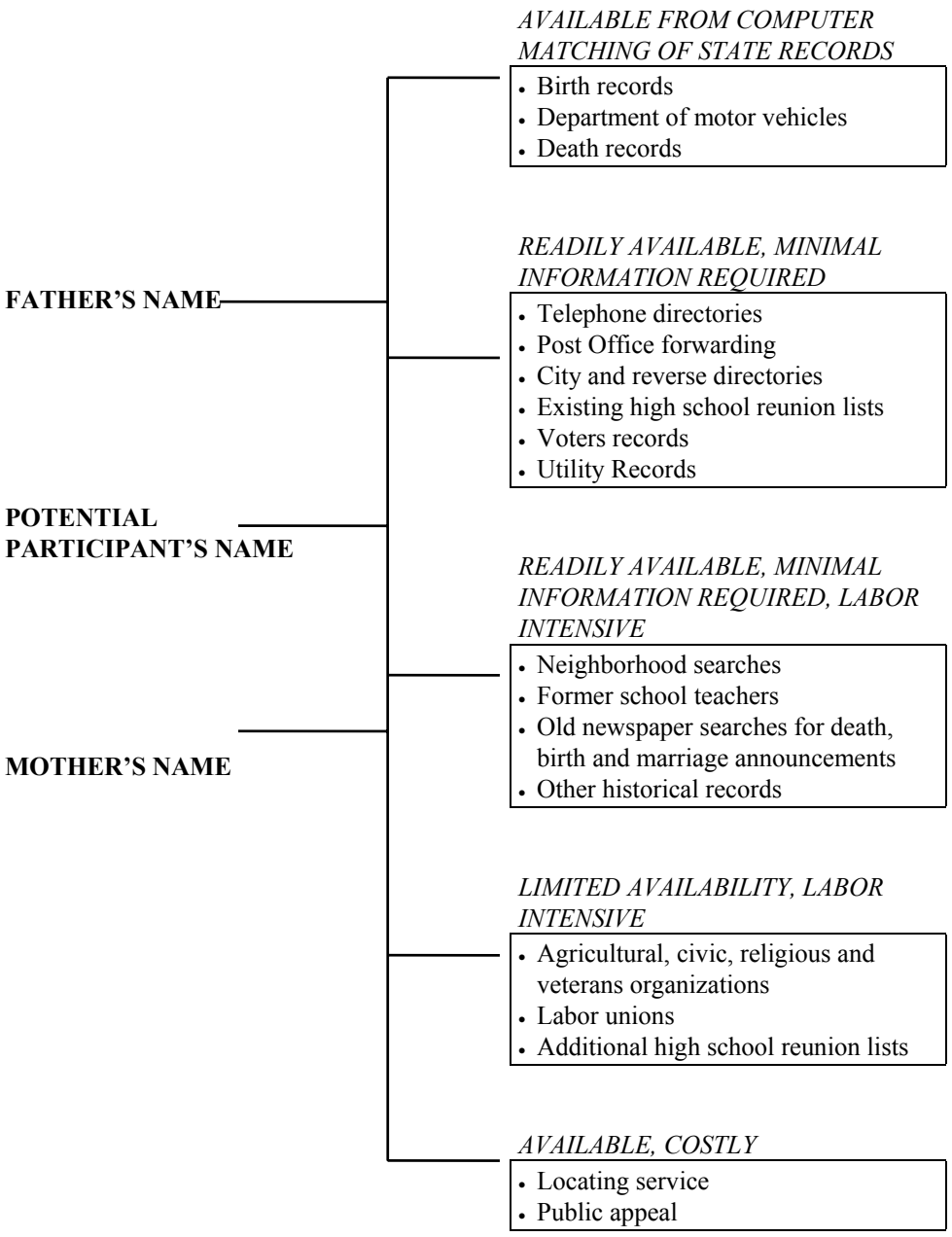
The results suggested that identifying, tracing, and locating residents of this region during the time period of interest regarding Hanford radiation releases was feasible, at least for the more rural segment of the population.

B.1.c. Overview of Tracing Efforts

At the time the HTDS protocol was written, tracing and locating large numbers of people who were born in the areas of interest up to 50 years ago presented significant logistical challenges. It was presumed this would be particularly true for women, many of whose surnames would have changed at least once in the intervening years because of marriage.

The initial approach to locating and tracing individuals is depicted graphically in Figure V.B-1. As shown in the diagram, searches were initiated based on the cohort member's name, the father's name and the mother's name. At the start of the search, efforts were concentrated on locating any one of these three individuals until enough information was obtained to focus on the location of the potential participant. The tracing process was to be undertaken as an investigative process, using different sources, as they needed to be pursued. The sources depicted were generally pursued in the order shown, beginning with the most readily available, least costly and least labor intensive, and progressing toward the most costly and most labor intensive until the potential participant was located or until all reasonable effort had been expended. See below for a more detailed description of the tracing effort.

Figure V.B-1. Locating and Tracing Potential Participants



Initially, two approaches were utilized to trace potential participants. The first was a series of computer matches of the study cohort to databases maintained by the State of Washington. These included: 1) death certificates; 2) recent birth certificates (second generation births), linking through both the father's and mother's name (i.e., the potential participant's name); and 3) Department of Licensing (Driver's License and Motor Vehicle Registration) information. The second approach was to use readily available and relatively inexpensive sources. These included primarily searches of telephone books, city directories, and Cole's reverse directories. In addition, several school reunion lists had already been obtained, and additional lists were sought as sources of potential participant follow-up. Field staff also searched for information using the Social Security Death Index kept by the Genealogical Library of the Church of Jesus Christ of Latter Day Saints**.** This source was particularly useful in locating family members through the deceased parent's Social Security payments.

Persons not located with the computer matches or through the use of readily available sources were sought using more intensive search methods including the use of county records of marriage records, and using local libraries to search old newspapers for obituaries, wedding and birth announcements. An attempt to locate individuals not found using any of the above sources was made through contacts in the communities.

To minimize the potential for bias in locating cohort members that might be related to either exposure or disease status, it was decided that several possible sources of information would not be used. These included population-based tumor registries, unsolicited self-report by members of the public, and mailing lists related to Hanford issues. Because each of these sources could be the only means of locating some persons, and inclusion in these sources might be related both to exposure status and/or thyroid disease status, they were specifically avoided for tracing purposes.

When contacting people who may have had information on the potential participant's location, such as the potential participant's mother or father, they were told that we were attempting to locate people who had been selected from birth certificates to participate in a medical research study. If the contact requested additional information, they were told that the study was the Hanford Thyroid Disease Study and were given basic information about the study. If the person requested additional information before disclosing the location of the potential participant, he or she was advised to contact the Seattle office toll-free telephone line. When contacting potential study participants a script was used to provide basic information about the HTDS and to inform them that a letter would be sent explaining the study in detail. If more information about the study was needed, the potential participant was advised to call the Seattle office toll-free telephone number and speak with the Participation Coordinator**.**

B.1.d. Staffing and Logistics

The tracing field staff consisted of several employees located in eastern Washington. A procedure manual was used to prioritize steps to be taken in locating potential participants. Records of each step in the search for each potential participant were kept to learn more about the most efficient methods for locating potential participants. Regular meetings were held in the Tri-Cities with the Project Manager and Principal Investigator to assess the success of this component throughout the fieldwork phase of the study.

B.2. Revisions to the Original Protocol Plan

B.2.a. Deletion of Ineffective Sources of Information and Addition of New Sources

During the Pilot Study, it was determined that some of the more difficult to use and often most expensive sources were less effective and these were not actively pursued in the Full Study. For example, high school reunion lists were moderately helpful in the Pilot Study (useful information was obtained in 41% of the cases in which such lists were used), but required rather extensive efforts to obtain. Overall, school records were not a major source of tracing information. The use of a locator service toward the end of the Pilot Study proved to be very expensive per potential participant located, and the results varied considerably. This source

was not routinely used beyond the Pilot Study, but was replaced by sources mentioned below, provided by the newly developed FHCRC Tracking Resource Center (TRC).

The use of more intensive contacts in small communities in the region was explored in the last few months of the Pilot Study Tracing process. Study staff made several trips during the Pilot Study to small towns to talk to local citizens and "old-timers" and to look through local records (e.g., marriage, utilities, property records). In general, these trips were moderately successful but labor intensive. Local postmasters, teachers, and community leaders were able to provide some guidance, and often helped to gain access to local records that might otherwise have been difficult to obtain. Such sources were used for only a very few potential participants, but when it was appropriate to pursue such sources, they were generally very successful in locating that potential participant.

In summary, the experience gained in the Pilot Study tracing effort identified a number of key approaches and sources of information that proved to be useful in locating potential participants. These approaches and sources defined the primary methodology used in locating the remaining potential participants needed for a Full Study. Those methods and sources that did not prove to be as useful were reserved for the most difficult to locate, when other resources had been exhausted.

B.2.b. Addition of Computer On-line Database Information

Late in the tracing process for the Full Study, a Tracking Resource Center (TRC) was developed by the FHCRC. This resource was designed to provide tracing, locating, and tracking services to a number of Center projects needing to identify and locate study participants or former patients. The TRC was utilized by the HTDS to locate potential participants who could not be located by other means. Additional new resources available through the TRC were: 1) a national, on-line database resource providing matches by name and previous address; and 2) a national, on-line database providing matches by name and date-of-birth, linking multiple public records available at that time. While the use of these resources tended to be more expensive, they were less labor intensive and frequently provided leads to assist in locating individuals who were the most difficult to locate.

B.3. Final Tracing Process

Tracing of potential participants was conducted in three stages, in the following chronological order for most potential participants. The first stage consisted of a series of linkages with publicly available data sources. The second stage, which constituted the majority of the tracing effort, utilized a variety of resources to look for potential participants on an individual basis. The third stage, undertaken only for those individuals most difficult to locate, was to enlist the services of the newly created FHCRC Tracking Resource Center and/or a professional locating company.

B.3.a. Linkages with Publicly Available Data Sources

Five types of linkages to publicly available data sources were performed on either the entire study sample or the appropriate subgroup, based on type of linkage. First, the study sample was manually matched to Washington State infant death certificates for the years 1942-1950 for the six original Pilot Study Counties

(not including Adams).[1] Second, the study sample was matched by computer to the Washington State Death Index (WSDI) for the years 1965-1990.[2] This included some records for Washington State residents who died outside of Washington, which were obtained through interstate exchange agreements. For females, matching to the death index was based on the potential participant's birth name from the birth certificate and the Father's surname as reported on the death certificate. Third, the Pilot Study sample was manually matched to a list of Washington State Vietnam War deaths. Because of the low return from searching the Vietnam War Deaths list (only one match was found for the entire Pilot Study sample), routine searching of this list was not continued for all potential participants in the Full Study samples, but was referred to as appropriate for more difficult to locate individuals.

The fourth step was perhaps the most unusual linkage undertaken in this series of linkages. Washington State birth certificates list mother's maiden (or birth) and current name, father's name, mother's and father's ages, and the child's name. To use this information to find study participants, female potential participant birth names were matched to mother's maiden name on Washington State birth certificates for the period 1956-1990 (second generation births), primarily to identify possible married names. For matches found in this way, the child's last name was assigned as a potential married name for the mother.

Fifth, the names of the entire study sample (including possible married names obtained in step four, above) were matched to the Washington State Department of Licensing (WSDOL) Driver's License Records by name and date of birth. This match was periodically re-run during the course of the study as the WSDOL records were updated. Matches were also conducted individually as new possible married names, children's names, and spouse names were identified.

After the final linkages for the entire group were performed and some potential participants located, three additional linkages were performed only on potential participants not yet located. For the Pilot Study, potential participant parents' names were matched by computer to the WSDI (1965-1990), using father's name, mother's maiden name and mother's potential married name from the potential participant's birth certificate. The purpose of this link was to provide dates and place of death for parents for whom an obituary could then be found. Because most obituaries list the survivors (including current names and place of residence), this information was sometimes used to locate the potential participant or a relative of the potential participant. Since the parents' dates of birth were not available on birth certificates (only age was listed), this linkage resulted in many possible matches and was not repeated for the Full Study Sample.

The second additional linkage performed for potential participants not located initially was to match males to father's name on Washington State birth certificates from 1956-1990 (second generation births) to identify possible children, through whom the potential participant may be located. Children identified this way were then matched to the WSDOL records.

The final computer linkage, completed only for the Pilot Study, was to match female potential participants' names (both birth and potential married names) to centralized Washington State marriage records, stored on microfiche, from 1970-present. New potential married names identified through this linkage were then matched to the WSDOL records. For the Full Study, this search was done on an individual basis, mainly at the county level, as the same access to these files was not possible at the time of the Full Study.

During the Full Study, matching against the National Death Index (NDI) files became available. The NDI is a central computerized index of death record information since 1979, compiled by the National Center of Health Statistics from information submitted by state vital statistics offices. Each record contains a standard set of identifying data for each decedent. Matches were performed for all potential participants not already located. In addition, parent names from the potential participant's birth certificate were matched for some potential participants. In this way, the informant listed on the parental death certificate could be used as a source for locating the potential participant.

[1] Infant death certificates were not routinely reviewed for 1940 and 1941 due to access problems.
[2] The WSDI was reviewed for all subjects classified as unable to locate throughout the tracing process, the most current issue of the WSDI was periodically reviewed for individuals not located in initial attempts.

Several scenarios were possible for information on a given potential participant from the data linkages. These included:

1. Data linkage shows potential participant deceased: For all potential participants linked to a Washington State Death Record, a death certificate (DC) was requested from the state. While these linkages were performed as "exact matches" (highest possible likelihood that this is the right person), great care was taken to check the death certificate against the birth certificate for any indication that this may be a mismatch.

2. Data linkage shows potential participant linked to subsequent birth: For female potential participants linked to subsequent births, the data from the match were checked against all information entered on the potential participant birth certificate and tracing sheet. If this appeared to be a likely match and a potential married name was elicited, tracing efforts were then directed toward this name until such time that it was confirmed this was indeed the right person.

3. Data linkage shows new address based on WSDOL: Linkage to this source was performed using birth certificate names as well as possible new names generated from the subsequent births listing. The new address information represented the current address held by the WSDOL for that person. The date the information was given to WSDOL as valid was included in the listing.

B.3.b. Manual Tracing Resources

Following the data linkages performed on the potential participant roster database, information was transferred to the eastern Washington staff to conduct tracing efforts manually for each individual. A copy of each potential participant's birth certificate was included in a file created for each potential participant which also included the tracing forms specifically designed for use in this study for documentation of tracing efforts undertaken (See Appendix 3).

The second stage of tracing activity utilized numerous sources identified and pursued by HTDS study staff. In general, after the initial linkages were complete and the matches provided, the tracing staff first undertook the process of locating those potential participants with the most promising information available. This approach was taken to ensure a steady supply of potential participants to be recruited and scheduled for clinics and to keep the study progressing as efficiently as possible. For example, potential participants (or their parents) with exact matches to the WSDOL data (providing addresses) were next searched through telephone directories, city/reverse directories and/or CD-ROM directories and other available resources.

Following the Pilot Study, several changes were made to the databases used to record tracing information. Information on the usefulness of sources was no longer collected. In addition, some sources, such as newspapers, were split into two or more categories to better capture the purpose of their use, such as locating obituary information. For this reason, separate tables are shown here for the Pilot Study Sample, and Transition and Full Study Samples. Tables V.B-1 and V.B-2 display the number of potential participants for whom each manual source was ever used for those in the Pilot Study Sample, and those in the Transition and Full Study Samples.

Overwhelmingly, the primary sources of information for tracing potential participants were telephone directories, family members, and various public records. Initially, phone book searches were conducted by hand, utilizing current and historical phone books obtained by the study and those available in local and regional libraries. Consequently, nearly all (97%) cohort members selected during the Pilot Study were sought in phone books (Table V.B-1). Early in the Pilot Study, however, CD-ROM products listing published phone numbers throughout the United States were acquired for this purpose and used extensively, but did not replace telephone directory use. About half of cohort members in the Transition and Full Study samples were sought in phone books, and 89% on CD-Rom directories (Table V.B-2). Directory assistance throughout the United States was also used extensively. Of the other types of sources used, the Social Security death rosters, City and County records (e.g., marriage records), and obituary information from newspapers and funeral homes were used the most. The category "Other Sources" in the following tables includes numerous other approaches

utilized by tracing staff that were limited to a very few potential participants for any given source. Table V.B-3 depicts the usefulness of sources in locating potential participants, based on the Pilot Study experience.

Table V.B-1. Tracing Sources Used, All Potential Participants – Pilot Study Sample

Source*	Source Was Ever Used (N=1587) †	
	No. of Potential Participants	Percent of Potential Participants (%)
Directories		
• Telephone directories (hard copy)	1546	97.4
• CD-ROM telephone directories	769	48.5
• Directory assistance	691	43.5
• City/Reverse directories	542	34.2
• CA People Finder/Western Gold	8	0.5
School Records		
• High school reunion lists	187	11.8
• Other school records	13	0.8
• Alumni organizations	8	0.5
• School registration records	4	0.3
• Former school teachers	2	0.1
Other Specific Sources		
• Relatives	850	53.6
• Social Security roster	689	43.4
• City/county records (includes marriage records)	307	19.3
• Locating service	189	11.9
• Online services*	130	8.2
• Death certificates	80	5.0
• Newspapers	63	4.0
• Funeral home/cemetery	56	3.5
• Neighborhood searches	51	3.2
• HTDS-ID letters	36	2.3
• Employers	35	2.2
• Libraries	28	1.8
• Postal service	17	1.1
• Veterans organizations	14	0.9
• Letter to Social Security Administration	12	0.8
• Native American sources (tribes and IHS)	9	0.6
• Other HTDS participants	9	0.6
• Birth certificates	8	0.5
• Religious organizations	6	0.4
• Civic organizations	6	0.4
• Agricultural organizations	4	0.3
• Utility records	2	0.1
• Labor unions	2	0.1
• Voter registration	2	0.1
• Military reunion lists	1	0.1
• Historical documents	1	0.1
Other sources	56	3.5

* The sources listed in this table do not include record linkages performed prior to entry of potential participants into the tracing system.
† Excludes 3 potential participants for whom no tracing data were entered due to a clerical error.

Table V.B-2. Tracing Sources Used, All Potential Participants – Transition and Full Study Samples

Source*	Source Was Ever Used (N=3475)†	
	No. of Potential Participants	Percent of Potential Participants (%)
Directories		
• CD-ROM telephone directories	3106	89.4
• Telephone directories (hard copy)	1750	50.4
• Directory assistance	1014	29.2
• City/reverse directories	919	26.4
• CA People Finder/Western Gold	28	0.8
School records		
• High school reunion lists	121	3.5
• School registration records	9	0.3
• Former school teachers	1	< 0.1
Other specific sources		
• Relatives	1440	41.4
• Social Security roster	1410	40.6
• Obituaries/funeral homes	1070	30.8
• Death index/death records	565	16.3
• Online services	491	14.1
• Marriage licenses	289	8.3
• Response to HTDS-ID letter	166	4.8
• Tax assessors	76	2.2
• Neighborhood searches	45	1.3
• Postal service	33	0.9
• Employers	32	0.9
• Locating service	22	0.6
• Letter to Social Security Administration	16	0.5
• List of Vietnam veterans	10	0.3
• Military locator service	10	0.3
• Voter registration	9	0.3
• Labor unions	3	0.1
• Other city/county records	3	0.1
• Civic organizations	2	0.1
• Agricultural organizations	2	0.1
• Religious organizations	2	0.1
• Veterans organizations	1	< 0.1
• Utility records	1	< 0.1
Other sources	90	2.6

* The sources listed in this table do not include record linkages performed prior to entry of potential participants into the tracing system.
† Table excludes n=134 potential participants who were not entered into the HTDS tracing system. All but one of these potential participants (n=133) were located with information obtained from record linkages with the Washington state Department of Licensing, prior to the implementation of the revised tracing system for transition and full study potential participants. As a result of clerical error, the remaining potential participant was never entered into the tracing system and, therefore, was not traced.

Table V.B-3. Usefulness of Tracing Sources in Locating Study Potential Participants – Pilot Study Only (N=1587)†

Source*	Source was Ever Used** No.	%	Source Lead to or Resulted in Locating Potential Participant No.	% ††
Directories				
• Telephone directories (hard copy)	1543	97.2	1165	75.5
• CD-ROM telephone directories	762	48.0	744	97.6
• Directory assistance	664	41.8	523	78.8
• City/reverse directories	535	33.7	240	44.9
School records				
• High school reunion lists	184	11.6	75	40.8
• Other school records	13	0.8	12	92.3
• Alumni organizations	8	0.5	7	87.5
• School registration records	4	0.3	2	50.0
• Former school teachers	2	0.1	0	--
Other specific sources				
• Relatives	820	51.7	812	99.0
• Social Security roster	672	42.3	312	46.4
• City/county records (includes marriage records)	304	19.2	154	50.7
• Locating service	187	11.8	85	45.5
• Newspapers	63	4.0	56	88.9
• Death certificates	51	3.2	44	86.3
• Neighborhood searches	50	3.2	32	64.0
• Funeral home/cemetery	35	2.2	28	80.0
• Employers	33	2.1	28	84.8
• Libraries	28	1.8	27	96.4
• Postal service	16	1.0	14	87.5
• Veterans organizations	14	0.9	7	50.0
• HTDS-ID letters	13	0.8	13	100.0
• Letter to Social Security Administration	12	0.8	12	100.0
• Native American sources (tribes and IHS)	9	0.6	7	77.8
• Other HTDS participants	9	0.6	9	100.0
• Birth certificates	8	0.5	8	100.0
• Religious organizations	6	0.4	6	100.0
• Civic organizations	6	0.4	5	83.3
• Agricultural organizations	4	0.3	4	100.0
• Utility records	2	0.1	1	50.0
• Labor unions	2	0.1	0	--
• Voter registration	1	0.1	1	100.0
• Military reunion lists	1	0.1	1	100.0
• Historical documents	1	0.1	1	100.0
Other sources	56	3.5	49	87.5

* The sources listed in this table do not include record linkages performed prior to entry of potential participants into the tracing system.
** These numbers differ from those in Table V.B-2 as this table contains only tracing performed prior to the end of the Pilot Study, after which information on usefulness of sources was no longer collected.
† Excludes 3 potential participants for whom no tracing data were entered due to a clerical error
†† Percent of those for who source was ever used.

B.3.c. Unlocated Potential Participants

Extensive efforts were made to locate each potential participant identified through birth certificates. While the tracing effort was extremely successful, not all potential participants could be located from the minimal information provided by a birth certificate from 50 years ago. Tables V.B-4 and V.B-5 show efforts expended on potential participants not located. Before efforts were closed out on any individual potential participant, an extensive amount of effort was required. This effort included all of the linkages performed in the initial tracing phase, along with four primary sources that would be tried for everyone. These sources included telephone directories, CD-ROM telephone directories, the Social Security death roster or WSDI, and one on-line service as mentioned above. These represented the only manual sources that were appropriate to try for all potential participants.

Table V.B-4. Tracing Efforts for Those Not Located – Pilot Study Sample

Source	Source Was Ever Used (N=78) No.	Percent of Unlocated (%)
Directories		
• Telephone directories (hard copy)	77	98.7
• CD-ROM telephone directories	77	98.7
• City/reverse directories	62	79.5
• Directory assistance	61	78.2
• CA People Finder/Western Gold**	5	6.4
School records		
• High school reunion lists	11	14.1
• School registration records	2	2.6
• Other school records	1	1.3
• Alumni organizations	1	1.3
Other specific sources		
• Social security roster	75	96.2
• Online services**	70	89.7
• Locating service	48	61.5
• City/county records (includes marriage records)	45	57.7
• Death certificates	21	26.9
• Relatives	11	14.1
• Neighborhood searches	5	6.4
• Funeral home/cemetery	5	6.4
• HTDS-ID letters	5	6.4
• Letter to Social Security Administration	4	5.1
• Newspapers	3	3.8
• Libraries	2	2.6
• Postal service	2	2.6
• Employers	1	1.3
• Native American sources (tribes and IHS)	1	1.3
• Other HTDS participants	1	1.3
• Birth certificates	1	1.3
• Agricultural organizations	1	1.3
• Labor unions	1	1.3
• Voter registration	1	1.3
Other sources	3	3.8

* The sources listed in this table do not include record linkages performed prior to entry of potential participants into the tracing system.
** Sources initiated during the transition/full study

Table V.B-5. Tracing Efforts for Those Not Located – Transition and Full Study Samples

Source	Source Was Ever Used (N=242) No.	Percent of Unlocated (%)
Directories		
• CD-ROM telephone directories	240	99.2
• Telephone directories (hard copy)	218	90.1
• City/reverse directories	155	64.0
• Directory assistance	122	50.4
• CA People Finder/Western Gold	26	10.7
School records		
• High school reunion lists	20	8.3
• School registration records	2	0.8
Other specific sources		
• Social security roster	232	95.9
• Online services	225	93.0
• Obituaries/funeral homes	158	65.3
• Death index/death records	139	57.4
• Marriage licenses	48	19.8
• Relatives	25	10.3
• Response to HTDS-ID letter	20	8.3
• Tax assessors	11	4.5
• Postal service	6	2.5
• Neighborhood searches	5	2.1
• List of Vietnam veterans	4	1.7
• Locating service	3	1.2
• Letter to Social Security Administration	3	1.2
• Military locator service	3	1.2
• Employers	2	0.8
• Voter registration	2	0.8
• Civic organizations	2	0.8
Other sources	13	5.4

* The sources listed in this table do not include record linkages performed prior to entry of potential participants into the tracing system.

Tracing potential participants based on birth certificates from the 1940s required the use of multiple sources of information in most cases. After the initial linkages were performed, more than one source of information was used for virtually all of the potential participants being traced. This is because even when a computer linkage was made (e.g., with WSDOL files), at least one additional step was almost always required to obtain information sufficiently detailed to contact the potential participant (e.g., a telephone number). For those most difficult to locate, many sources may have been used before the potential participant was located or determined to be unlocatable. Tables V.B-6 and V.B-7 summarize the extent to which multiple sources were used in the Pilot Study Phase and the Transition and Full Study Phases, by tracing outcome. The number of sources used after the initial linkages ranged from 1 - 16, with a mean of 4.8 sources per potential participant for the Transition and Full Study Phases.

Table V.B-6. Number of Sources Used to Trace Located and Unlocated Individuals – Pilot Study Sample

Number of Sources Used	Located		Unlocated		Total	
	No.	%	No.	%	No.	%
1	13	0.9	0	--	13	0.8
2	264	17.5	0	--	264	16.6
3	242	16.0	0	--	242	15.2
4	214	14.2	1	1.3	215	13.5
5	206	13.7	3	3.8	209	13.2
6	188	12.5	8	10.3	196	12.4
7	135	8.9	16	20.5	151	9.5
8	111	7.4	13	16.7	124	7.8
9	50	3.3	13	16.7	63	4.0
10	37	2.5	10	12.8	47	3.0
11	28	1.9	6	7.7	34	2.1
12	9	0.6	5	6.4	14	0.9
13	4	0.3	2	2.6	6	0.4
14	3	0.2	1	1.3	4	0.3
15	3	0.2	0	--	3	0.2
16	0	--	0	--	0	--
17	2	0.1	0	--	2	0.1
Total	1509	100.0	78	00.0	1587†	100.0
Mean no. of sources	4.99		8.54		5.16	

* The sources listed in this table do not include record linkages performed prior to entry of potential participants into the tracing system.
† Table excludes 3 potential participants for whom no tracing data were entered due to a clerical error.

Table V.B-7. Number of Sources Used to Trace Located and Unlocated Individuals – Transition and Full Study Samples

	Located		Unlocated		Total	
No. of Sources Used	No.	%	No.	%	No.	%
1	108	3.3	0	--	108	3.1
2	474	14.7	2	0.8	476	13.7
3	664	20.5	1	0.4	665	19.1
4	563	17.4	0	--	563	16.2
5	334	10.3	11	4.5	345	9.9
6	307	9.5	19	7.9	326	9.4
7	240	7.4	38	15.7	278	8.0
8	210	6.5	46	19.0	256	7.4
9	149	4.6	48	19.8	197	5.7
10	93	2.9	30	12.4	123	3.5
11	50	1.5	23	9.5	73	2.1
12	22	0.7	12	5.0	34	1.0
13	9	0.3	5	2.1	14	0.4
14	7	0.2	5	2.1	12	0.3
15	2	0.1	2	0.8	4	0.1
16	1	< 0.1	0	--	1	< 0.1
Total	3233	100.0	242	100.0	3475†	100.0
Mean no. of sources	4.80		8.69		5.07	

* The sources listed in this table do not include record linkages performed prior to entry of potential participants into the tracing system.
† Table excludes n=134 potential participants who were not entered into the HTDS tracing system. All but one of these potential participants (n=133) were located with information obtained from record linkages with the Washington state Department of Licensing, prior to the implementation of the revised tracing system for transition and full study potential participants. As a result of clerical error, the remaining potential participant was never entered into the tracing system and, therefore, was not traced.

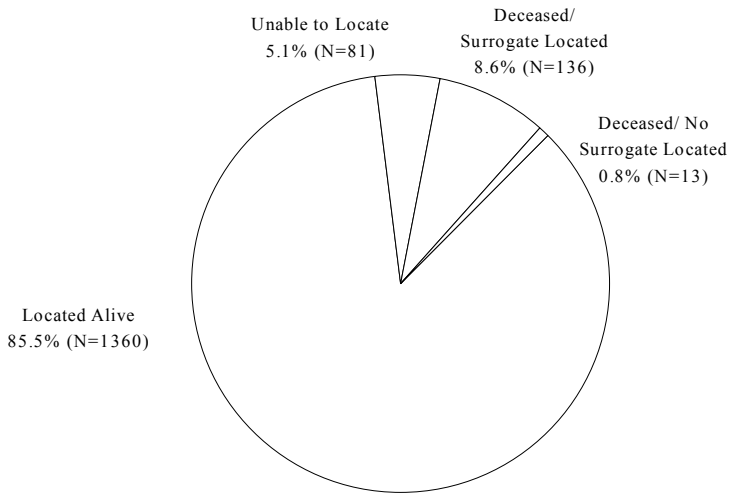
The effort expended toward the location of each potential participant was reviewed by the Data Collection Specialist (DCS) responsible for the case and by the Lead Data Collection Specialist or the Participation Supervisor to assure that all reasonable effort had been made and that all appropriate resources had been used to locate the individual. Only after this review was any potential participant "retired" as unable to locate. At least five sources were used for 316 of the 320 unlocated participants, and an average of more than eight sources were used before potential participants were "retired."

B.4. Outcome and Final Results

B.4.a. Results from the Pilot Study Sample

Figure V.B-2 displays the results of the tracing of individuals selected for the Pilot Study Sample. It should be noted that these numbers vary slightly from the Pilot Study Report of January 24, 1995, as additional potential participants from the Pilot Study selection were included in the Full Study. Of the 1590 individuals selected for the Pilot Study Sample, 1360 living individuals were located and 149 individuals were confirmed deceased. For 136 of the deceased individuals, a surrogate (ie. someone who might be able to provide study information about the deceased individual) was located. Thus, 94.9% of the original sample were located. Only 81 (5.1%) potential participants were listed as "unable to locate."

Figure V.B-2. Tracing Outcome for Pilot Study Sample (N = 1590)



B.4.a.1. Results by Strata

Figure V.B-3 presents tracing outcomes separately for the 791 females and the 799 males in the Pilot Study Sample. Success in locating living individuals was approximately the same for both sexes (87.2% for females and 83.9% for males). However, a larger proportion of Pilot Study Sample males were confirmed deceased (11.8%) than females (7.0%). Thus, after combining the living potential participants with the confirmed deceased, 95.6% of the Pilot Study Sample males and 94.2% of the females were located. This finding was somewhat unexpected, as it was anticipated that females would be much more difficult to locate after such a long period of time, due to name changes with marriage, particularly among this age group.

Figure V.B-3. Tracing Outcome for HTDS Pilot Study Sample, by Sex (N = 1590)

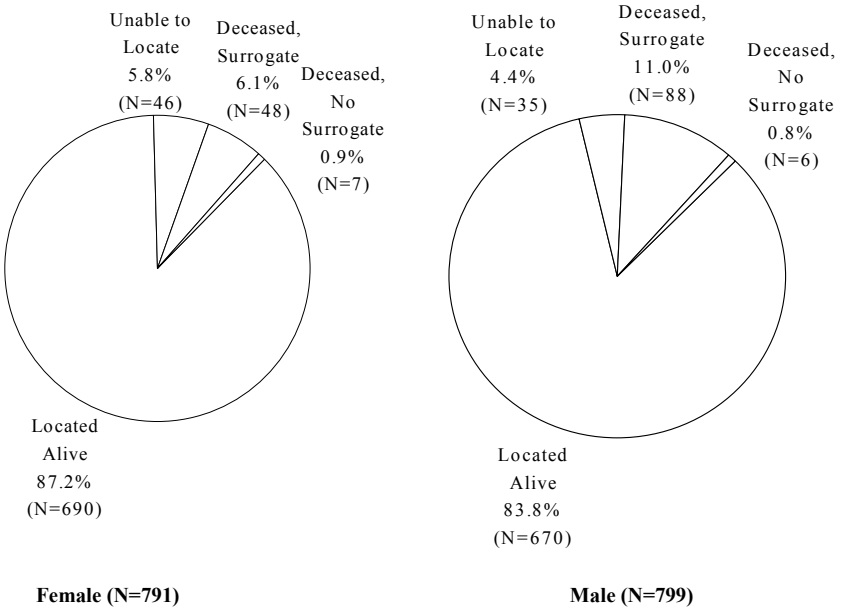


Figure V.B-4 displays the percent of the Pilot Study Sample located according to year of birth, from 1942-1946. There was relatively little difference in the proportion located in each birth year (range = 92.8% to 97.4%). Persons born in 1942 were slightly more frequently located than those born in the other years, but not substantially so.

Figure V.B-4. Tracing Outcome for Pilot Study Sample, by Year of Birth (N = 1590)

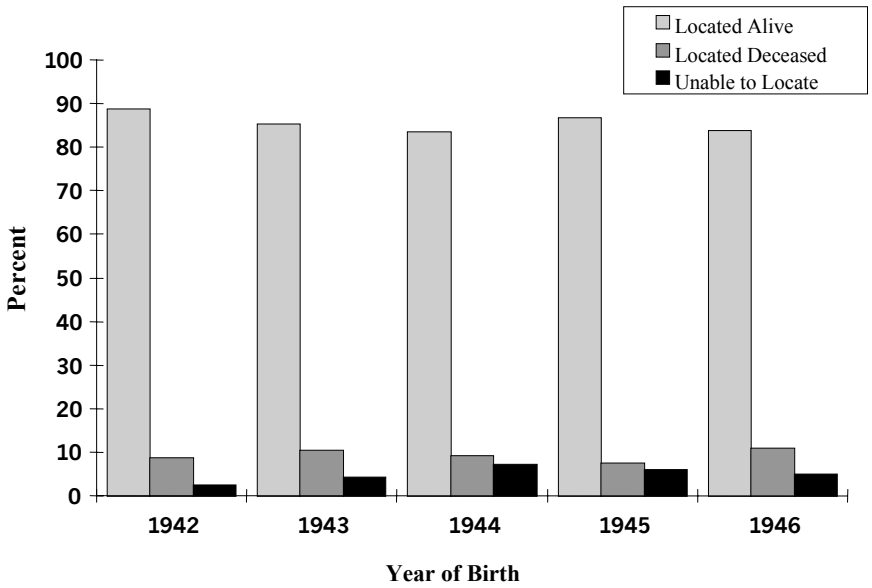


Figure V.B-5 displays the percent of the Pilot Study Sample located according to the eight regions that defined the geographic sampling strata (geostrata). These are arranged in the figure to correspond in an approximate manner to more urbanized areas (Richland, Pasco/Kennewick, and Walla Walla city) and predominately rural areas (counties, outside city). Although there is relatively little difference in the proportion located across the eight regions, there was a tendency for the more rural areas to have higher success rates. Such a pattern might be expected given that it is likely that the temporary workers who came to the area for construction jobs at the Hanford facility lived in more urban areas. In four of the five rural geostrata, 95% or more of the potential participants were located. Location rates for the three urban geostrata areas ranged from 89.3% to 93.6%.

Figure V.B-5. Tracing Outcome for Pilot Study Sample, by Geostratum (N = 1590)

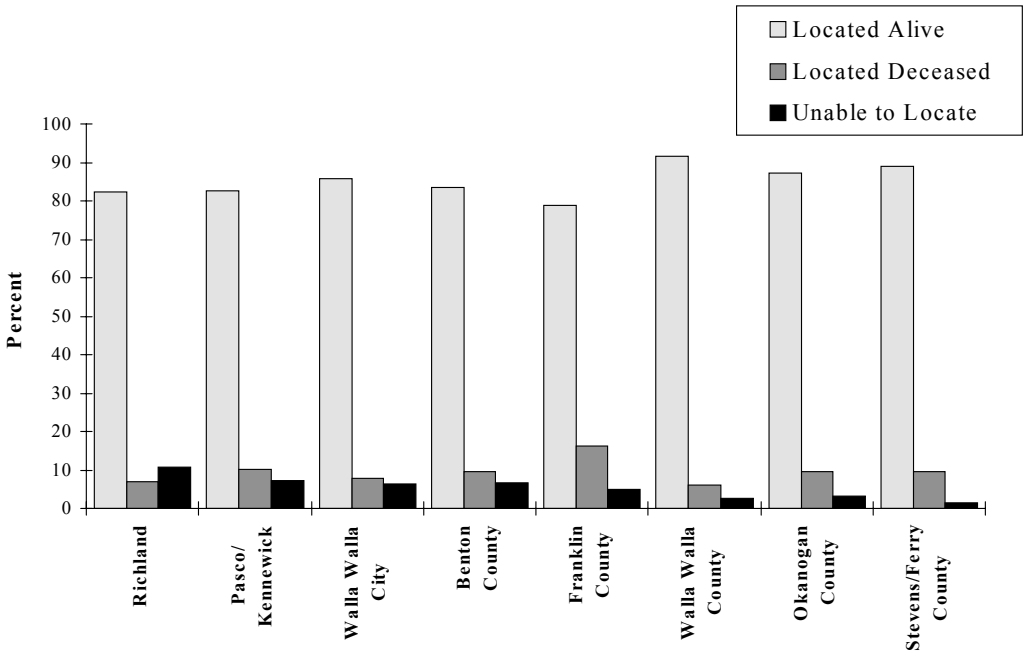


Table V.B-8 shows the proportion of Pilot Study Sample members located within each of the 76 strata (defined by gender, year of birth, and geostrata). More than half the sample was located in all 76 strata. At least 90% of the potential participants were located in 69 (90.8%) of the strata and at least 80% were located in 75 (98.7%) of the strata. In one stratum, 1944 Richland females, only 18 of 23 (78.3%) were located.

Table V.B-8. Percentage of Pilot Study Potential Participants Located in Each of the 76 Sampling Strata -Pilot Study Sample (N= 1590)

Geostratum and Sex**	Year of Birth 1942 No.	1942 %	1943 No.	1943 %	1944 No.	1944 %	1945 No.	1945 %	1946 No.	1946 %
Richland										
Female	See*		See*		18	78	19	90	18	86
Male	Footnote*		Footnote*		21	91	21	100	20	91
Pasco/Kennewick										
Female	21	100	18	90	21	95	16	80	18	90
Male	20	95	19	90	18	90	22	100	19	95
Walla Walla (city)										
Female	22	100	18	90	18	90	17	85	19	95
Male	19	95	20	100	19	90	20	95	20	95
Benton County										
Female	26	100	19	90	18	82	17	85	24	100
Male	22	92	19	90	20	95	22	100	24	96
Franklin County										
Female	21	95	12	100	22	96	7	88	16	100
Male	16	94	13	100	17	94	9	90	19	90
Walla Walla County										
Female	24	100	23	100	22	100	24	100	19	90
Male	26	96	22	100	23	96	26	96	19	95
Okanogan County										
Female	20	95	21	100	20	95	19	90	21	100
Male	22	100	19	95	21	95	20	95	21	100
Ferry/Stevens Counties										
Female	21	100	22	100	22	100	21	100	21	95
Male	21	100	20	95	21	95	24	100	20	100

* Richland was defined as a geostratum separate from Benton County only for births in 1944-1946.

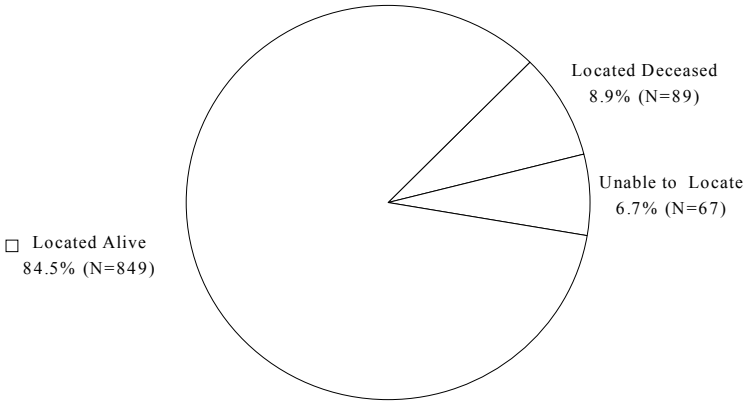
** Sex is defined as actual sex, not sex strata, as 2 potential participants were misclassified: one potential participant in the 1945 Pasco/Kennewick stratum was misclasssified as female and another potential participant in the 1946 Franklin County stratum was misclassified as male.

In summary, the Pilot Study demonstrated the feasibility of locating cohort members identified from birth certificate records from the early to mid-1940s. Overall, 91% of the 1590 Pilot Study Sample members identified from birth certificates were located by the end of the Pilot Study in December 1995. By the end of the Full Study, this percentage rose to 94.9%. Success in locating people did not differ substantially according to sex, year of birth, or geographic area of birth. This indicated that the methods used throughout the Pilot Study were effective at locating a substantial percentage of all selected potential participants.

B.4.b. Results from Transition Sample

The Transition Sample selection included potential participants from five of the eight geostrata used during the Pilot Study (Richland, Pasco/Kennewick, Walla Walla City, Benton County outside of Pasco, Walla Walla County outside of Walla Walla City). No further selections were made from the Okanogan or Ferry/Stevens geostrata because it was anticipated that the design of the Full Study would limit further selections to only geostrata near Hanford. No further selection from the Franklin County geostratum was possible, since all its members were selected for the Pilot Study. Figure V.B-6 shows the tracing outcomes for potential participants in the Transition Sample. A slightly lower percentage of potential participants were located in the Transition Sample (93.3%) than the Pilot Study Sample (94.9%). (Note that the category, "Deceased, Surrogate Located" was used only in the Pilot Study, as it was initially intended that deceased potential participants would be represented in the study by a surrogate, when available. Please see section V.D for a complete discussion of this issue.)

Figure V.B-6. Tracing Outcome for the Transition Sample (N = 1005)



B.4.b.1. Results by Strata

Figure V.B-7 shows tracing outcomes by sex for the Transition Sample. Tracing efforts were slightly more successful in locating males than females. This difference was more evident in the Transition Sample (95.0% vs. 91.6% of males and females respectively) than in the Pilot Study Sample (95.6% vs. 94.2%). However, the overall success rate for tracing both males and females remained consistently high.

Figure V.B-7. Tracing Outcome for Transition Sample, by Sex (N = 1005)

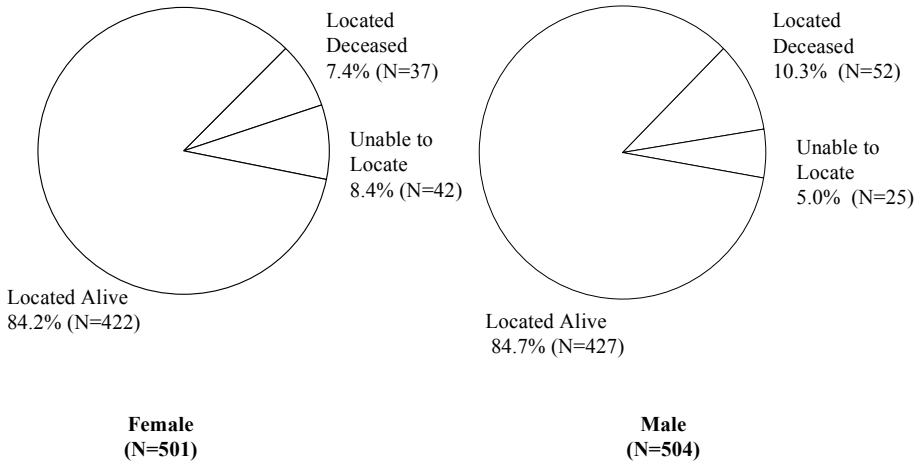


Figure V.B-8 shows tracing outcome by year of birth for the Transition Sample. Ability to locate potential participants in the Transition Sample ranged from 90.8% for those born in 1945 to 97% for those born in 1943.

Figure V.B-8. Tracing Outcome for Transition Sample, by Year of Birth (N = 1005)

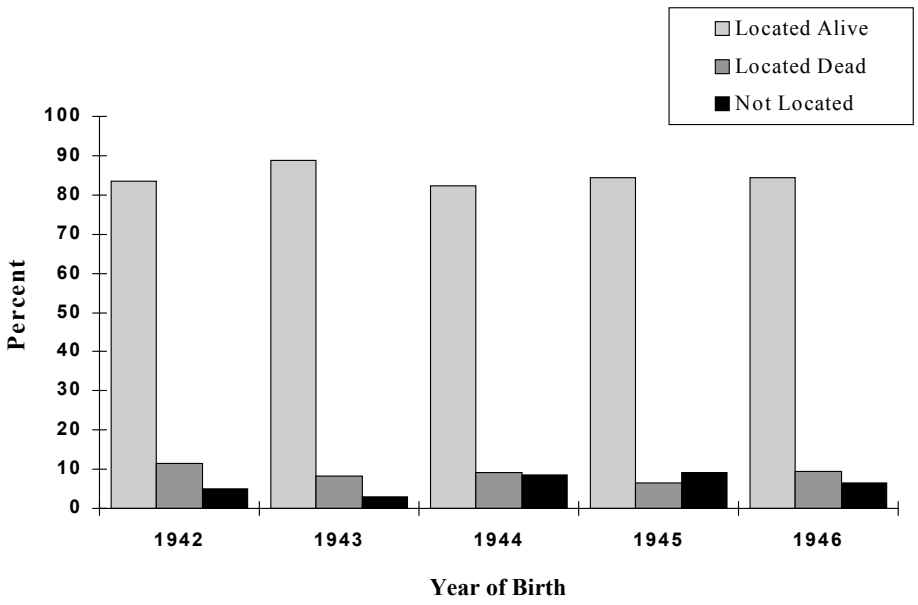
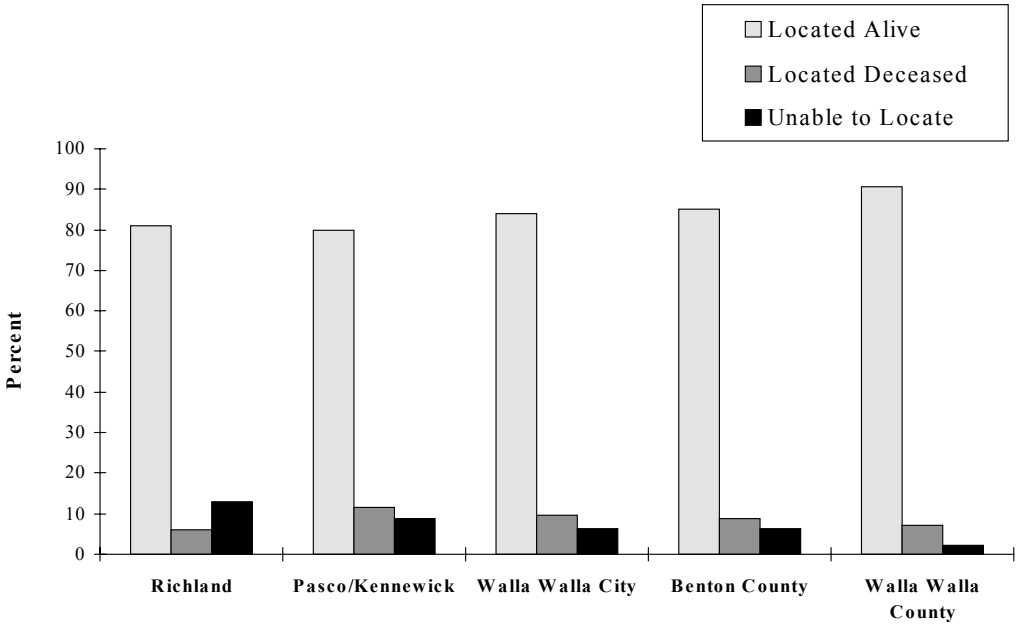


Figure V.B-9 shows tracing outcomes by geostrata for the Transition Sample. The success rate for locating Transition Sample members ranged from 87.1% in the Richland geostratum to 97.8% in the Walla Walla County geostratum. As in the Pilot Study Sample, efforts were slightly more effective in locating those born in rural areas, presumably due to a less mobile population.

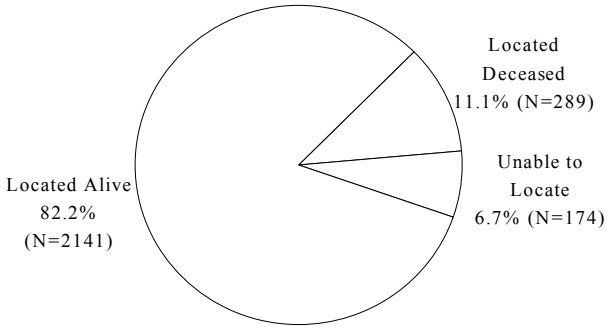
Figure V.B-9. Tracing Outcome for Transition Sample, by Geostratum (N = 1005)



B.4.c. Results for the Full Study Sample

The HTDS Full Study included all potential participants selected for the Pilot Study and Transition Samples, along with those selected later. For convenience, however, those selected after the Pilot Study and Transition Sample selections are designated the Full Study Sample. A total of 2604 potential participants were included in the Full Study Sample. The Full Study Sample was selected from five of the nine geostrata used in the entire study (Richland, Pasco/Kennewick, Benton County outside of Pasco, Franklin County outside of Kennewick, and Adams County). In addition, the Full Study Sample included people born in 1940-41. Figure V.B-10 shows tracing outcomes for the Full Study Sample. A larger percentage of Full Study Sample members were located deceased (11.1%, compared to 9.4% and 8.9% for the Pilot Study and Transition Samples), presumably due to the inclusion of potential participants born in 1940 and 1941. Nevertheless, the percentage located (93.3%) was similar to those for the Pilot Study and Transition Samples.

Figure V.B-10. Tracing Outcome for the Full Study Sample (N = 2604)



B.4.c.1. Results by Strata

Figure V.B-11 shows tracing outcomes by sex for the Full Study Sample. The percentage of female potential participants located in the Full Study Sample (91.3%) was again slightly lower when compared to the percentage of males located (95.2%). While the difference was slightly larger than for the Pilot and Transition samples, it was not substantially different.

Figure V.B-11. Tracing Outcome for the Full Study Sample, by Sex (N = 2604)

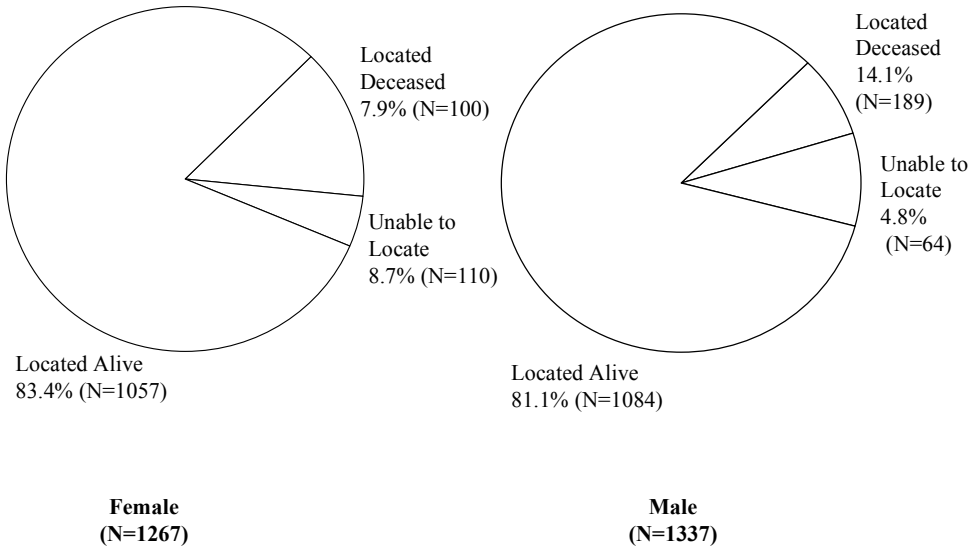


Figure V.B-12 shows tracing outcome by year of birth for the Full Study Sample. Ability to locate potential participants by year of birth in the Full Study Sample varied from 90.0% for those born in 1945 to 98.9% for those born in 1942. This is consistent with the location rates for the Pilot and Transition Samples.

Figure V.B-12. Tracing Outcome for the Full Study Sample, by Year of Birth

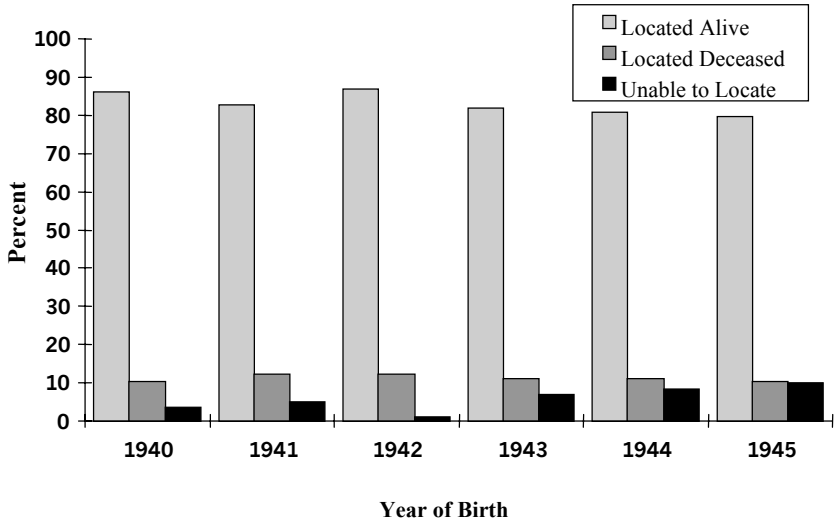
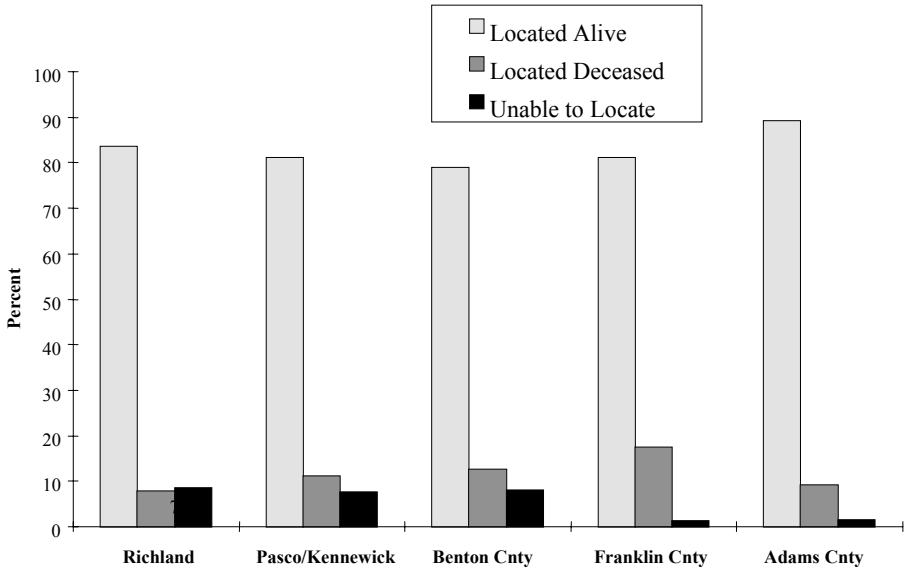


Figure V.B-13 shows tracing outcome by geostratum for the Full Study Sample. Success in locating potential participants ranged from 91.4% in the Richland geostratum to 98.6% in the Franklin County geostratum. The highest success was achieved in locating potential participants born in the relatively rural Franklin County and Adams County geostrata.

Figure V.B-13. Tracing Outcome for the Full Study Sample, by Geostratum



B.4.d. Overall Results for the Full Study

B.4.d.1. Success in Locating Study Potential Participants

Figure V.B-14 shows the final tracing outcomes for the entire study. Of the 5199 individuals sought, 4350 (83.7%) living individuals were located and 527 (10.1%) individuals were confirmed deceased. Thus, 93.8% of the sample were located and their identities confirmed. Only 322 potential participants (6.2%) remained unlocated at the end of the study. In addition, the ability to locate well over 90% of all potential participants did not vary substantially by sex, geographic region at birth, or year of birth. Figures V.B-15 to V.B-17 show the final tracing outcomes for the study by sex, by year of birth, and by geostrata.

Figure V.B-14. Final Tracing Outcomes for Entire Study (N=5199)

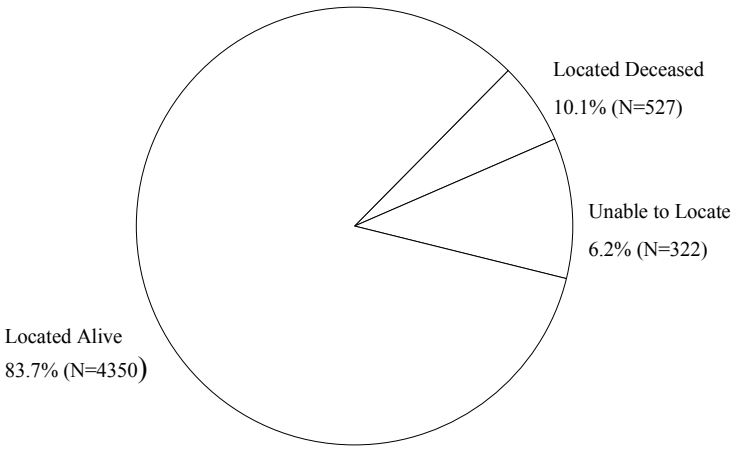


Figure V.B-15. Final Tracing Outcome for Entire Study, by Sex (N=5199)

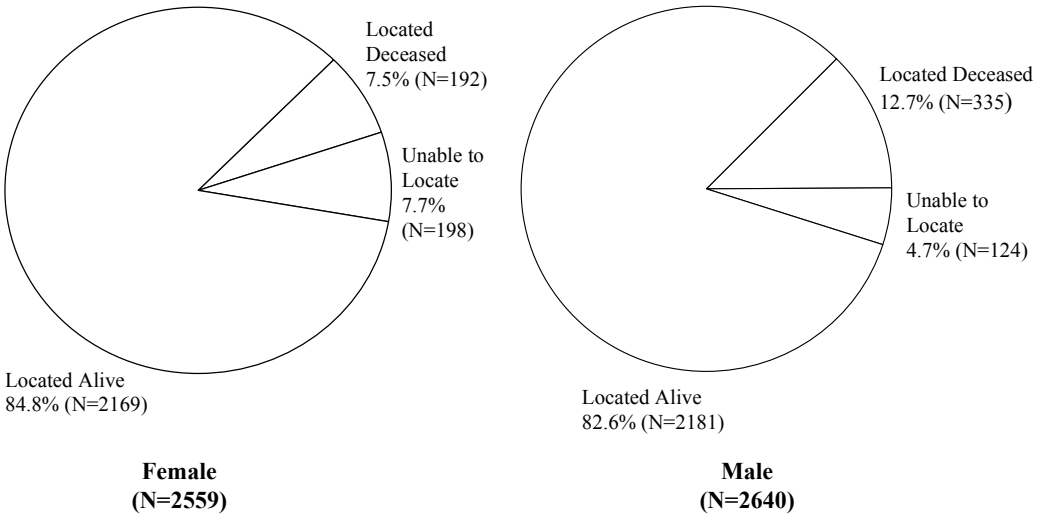


Figure V.B-16. Final Tracing Outcome for Entire Study, by Year of Birth (N=5199)

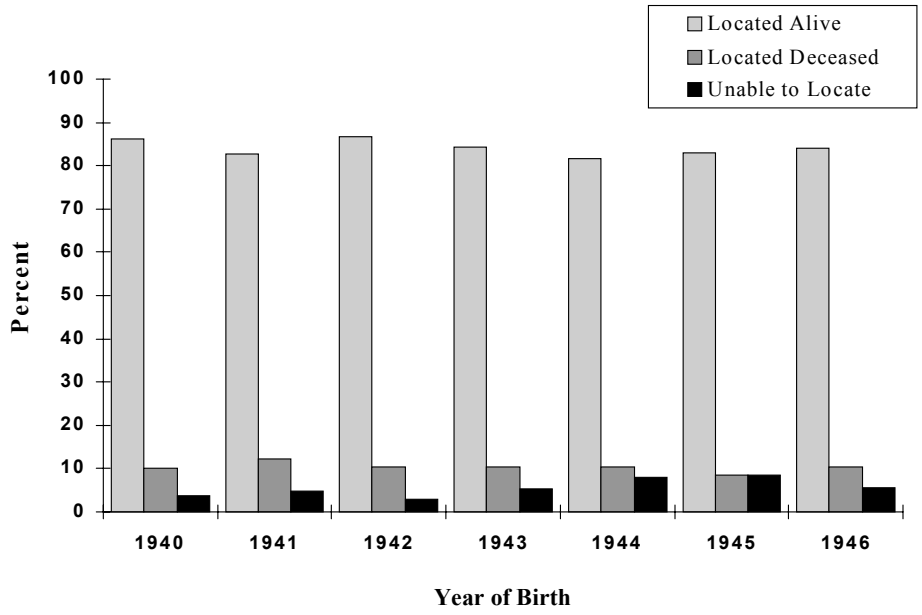
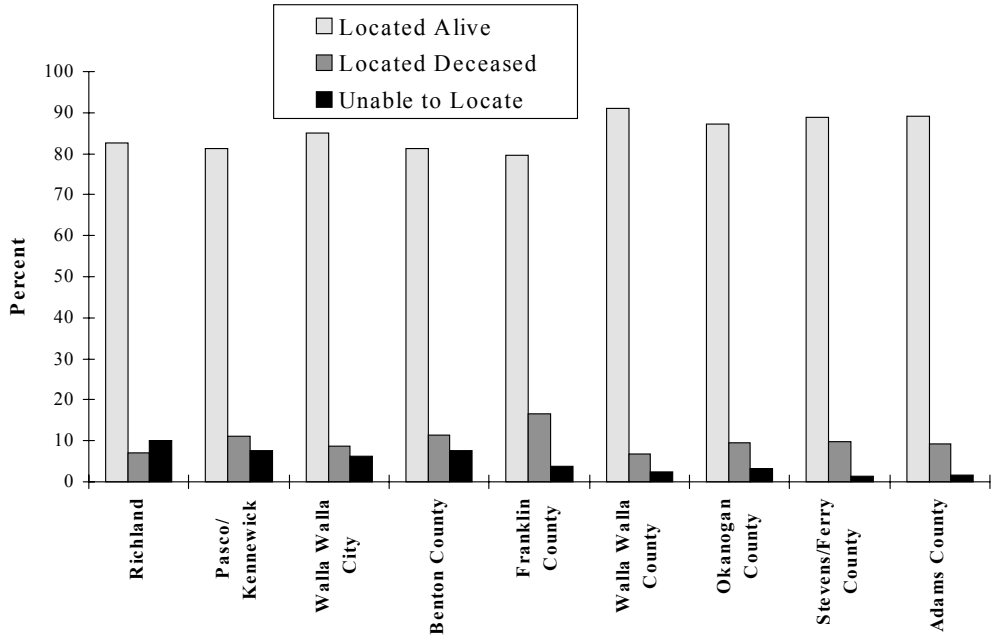


Figure V.B-17. Final Tracing Outcome for Entire Study, by Geostratum (N = 5199)



Almost 84% of all potential participants were located as living and potentially evaluable (whether they agreed to participate or not). For most (83.4%) this was confirmed directly by contact with the potential participant or with a close relative who could verify the potential participant's identity and provide a current

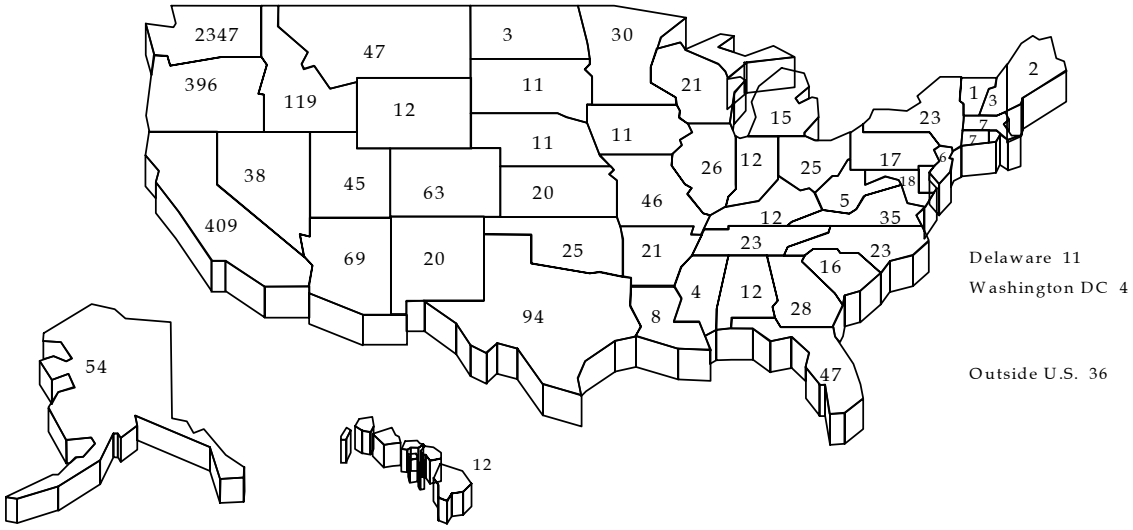
address. An additional 12 potential participants (0.2%) were located to a current address using other reliable sources providing enough information to verify their identity as the selected potential participant. These potential participants were sent a letter asking them to participate, but no direct contact was made, as no telephone number was available.

Five hundred twenty-seven (10.1%) of all selected potential participants were confirmed as deceased by a close relative and/or other reliable source (e.g., death certificate). A larger proportion of males was confirmed deceased (12.7%) than females (7.5%). Sixteen potential participants (0.3%) were located as living, but died during the study before agreeing to participate or prior to attending a clinic. An additional 22 (0.4%) potential participants were located as living but required a surrogate due to mental or physical conditions. Twenty-one of these potential participants were unable to participate in clinics, but were positively identified, while one attended a clinic with the assistance of a family member and an assistant. Two potential participants were determined to be ineligible, one because he was adopted and actually had been born in Spokane, which was his birth mother's usual residence, and the other due to a mistake on the birth certificate regarding the "Mother's Mailing Address," which indicated the mother lived in Walla Walla when the potential participant was born, when it should have been Columbia County as reflected in the "Mother's Usual Residence."

B.4.d.2. Current Residence of Living Potential Participants

At least one potential participant was located in every state in the U.S. except for Rhode Island (see Figure V.B-18). Fifty-four percent of the located potential participants resided in Washington State, 9.4% in California, 9.1% in Oregon and 2.7% in Idaho. The only other state where more than 2% of the located potential participants resided was Texas (2.2%). Thirty-six potential participants (0.8% of those located) resided outside of the U.S. Potential participants were located in Canada, Dubai, Ecuador, Germany, Mexico, Saudi Arabia, South Africa, England, Guam, Australia, Japan, France, Saipan, Hungary, Columbia, Taiwan and South Korea. Many of these (26) participated in the study. Although excessive travel costs associated with foreign travel prohibited the study from flying participants to the U.S. from outside North America, those participants who had plans to be in the U.S. during the study were brought to Seattle whenever possible to attend a clinic during that time.

Figure V.B-18. Current Residences of Located Potential Participants



B.4.d.3. Death Certificates Obtained for Deceased Potential Participants

As noted above, 527 cohort members were deceased when located, and another 16 located individuals died before agreeing to participate or attending a study clinic. For these 543 potential participants, an attempt was made to obtain a death certificate to verify the death and collect information about the cause of death. In 504 cases (92.8%), the death certificate was obtained. However, in 7.2% of cases, no death certificate could be located in the state in which the potential participant was reported to have died, or no state of death was known by the respondent, and could not be ascertained from the National Death Index. Consistent with the tracing results for living potential participants, the majority of those deceased had died in Washington State, with Oregon and California having the second and third largest proportions of deceased potential participants. Table V.B-9 shows the success in obtaining death certificates in the states where the 543 potential participants were reported to be deceased.

Table V.B-9. Summary of Death Certificates Obtained for Deceased Study Potential Participants

Reported Residence at Death	Death Certificates Obtained	Death Certificate Not Requested Due to Lack of Information	Death Certificate Requested But Not Found	Total
Washington	372	3	14	389
Oregon	30	0	1	31
California	22	0	5	27
Idaho	7	0	3	10
Montana	7	0	1	8
Texas	7	0	0	7
Colorado	6	0	0	6
Minnesota	6	0	0	6
Utah	5	0	0	5
Alaska	4	0	0	4
Nebraska	3	0	1	4
New York	3	0	1	4
Wyoming	2	0	2	4
Arizona	2	0	2	4
Nevada	3	0	0	3
Arkansas	2	0	0	2
Delaware	1	0	1	2
Florida	2	0	0	2
Georgia	2	0	0	2
Hawaii	2	0	0	2
Pennsylvania	2	0	0	2
South Carolina	2	0	0	2
Tennessee	1	0	1	2
Alabama	1	0	0	1
Illinois	1	0	0	1
Kentucky	1	0	0	1
Missouri	1	0	0	1
Mississippi	1	0	0	1
North Carolina	1	0	0	1
New Hampshire	1	0	0	1
Ohio	1	0	0	1
Oklahoma	1	0	0	1
Virginia	1	0	0	1
Out of U.S.	1	0	1	2
Unknown	0	3	0	3
Total	504	6	33	543

C. Recruiting

C.1. Background

The design of the HTDS posed significant challenges for recruiting study participants. It required that each participant be asked to identify an older relative (preferably the mother) to complete an extensive telephone interview. In addition, the participant was asked to travel to a clinic for a complete medical evaluation to determine the presence of thyroid disease and to complete an In-Person Interview.

C.1.a. Objectives of Recruiting

The objectives of the recruiting activities were to contact and obtain agreement of living potential participants to participate in the study, and to identify an appropriate Computer Assisted Telephone Interview (CATI) respondent. It was necessary to do this within a time frame that would provide sufficiently large pools of participants to schedule regional clinics, while also allowing participants ample opportunity to attend a clinic.

C.2. Recruiting Procedures

C.2.a. Initial Written Contact and Attempt to Contact by Phone

Each potential participant who was located was sent an introduction/participation letter, a study Fact Sheet and Description of Study Participation. Calls to potential participants began approximately 5-7 days after the letters were sent. Since potential participants would be in their late 40s and early 50s at the time of recruitment, they were assumed to be working. Therefore, the majority of recruitment calls were made during the evenings, taking into consideration the time zone in which the potential participant resided. A minimum of 10-15 evening attempts were made at various weeknight and weekend (generally Sunday evening) time periods, and a minimum of three daytime (weekend and weekday) calls were attempted.

If the potential participant could not be contacted by phone after 20-25 attempts, a second letter was sent explaining that the Recruiter had been unable to reach them at that phone number, and asking them to contact the Recruiter or Participation Coordinator at the toll-free HTDS number[1]. Further attempts to contact these people were postponed for approximately one month, after which both day and evening attempts would begin again until the participant was reached or until 40 attempts had been made. After 40-45 attempts resulting in no contact with either the potential participant or a household member, another letter was sent. This letter included the toll-free number and card for the potential participant to complete and return in a self-addressed stamped envelope confirming that he or she was the identified person, and asking for a phone number and time at which he or she would most likely be available.

If there was no response to this letter, and the letter was not returned with an address correction or as unable to deliver within one month, the potential participant was considered "unable to contact" and no further attempts were made. If the potential participant or household member had been reached at least once, additional attempts beyond the 40-45 calls were sometimes made, dependent on the nature of the contact.

If at any point in the process of attempting to contact a potential participant the phone number was disconnected or proved to be a wrong number, efforts were made to obtain a correct phone number through directory assistance, the original informant, or another available source. When necessary, potential

[1] If all or most calls up to this point resulted in an answering machine, a general message was left on up to three occasions explaining that we were calling from the HTDS and leaving the toll-free number for the potential participant to contact the Recruiter.

participants were returned to the tracing staff for additional tracing effort. Later in the recruitment phase, the FHCRC Tracking Resource Center assisted in locating updated address and telephone information. If a new telephone number could not be obtained, a second letter (as described above for those not contacted after 20-25 attempts) was sent asking the potential participant to contact HTDS at the toll-free number. If no response was received and no telephone contact was ever made, a letter was sent, asking him or her to contact the study or return the enclosed card to confirm identity as the study potential participant and/or providing contact information.

C.2.b. Telephone Contact with the Potential Participant

Once the potential participant was reached by phone, the Recruiter explained the reason for the call, confirmed the person's identity (full name, date and county of birth) and attempted to obtain agreement to participate by explaining the purpose and nature of the study and responding to any concerns as appropriate. A script was developed and used to ensure that each potential participant received the same basic information about the study. However, the nature of the recruitment call required that Recruiters be flexible enough to respond to individual questions and concerns. Recruiters were trained to be able to address a variety of questions and concerns in order to obtain the highest participation rate possible.

Every effort was made to recruit and provide interpreters or other assistance as necessary for non-English speakers, illiterate, hearing and vision impaired, or otherwise impaired persons to achieve maximum participation in the study. If a potential participant was reluctant to participate, the Recruiter would encourage him or her to contact the Participation Coordinator or Project Manager so that specific concerns could be addressed. If the potential participant still refused, in most cases, a second contact by letter was sent in approximately three to six months to allow for the possibility that he or she would have reconsidered the decision. This letter was followed by a telephone call, as with the first attempt, to try again to gain participation and to respond to any questions or concerns.

If the potential participant agreed to participate in the study, the Recruiter requested that the participant name a respondent for the CATI. Whenever possible, the CATI respondent was the potential participant's mother, who was assumed to be best able to answer the CATI questions. The next choices, in order of preference were: father, older sibling (at least six years older), or other family member who lived with the potential participant for a large part of his/her early childhood since birth.

The Recruiter described the dosimetry materials and interview to ensure that the participant felt that the respondent would be willing and able to complete the CATI process. A CATI Respondent Assessment (Appendix 4) was then completed with the participant. Questions were asked about the respondent's abilities, such as sight, hearing and special needs. This information was then provided to the CATI Interviewer.

If no CATI respondent was available, the potential participant was informed that he or she was still eligible for participation and was assured that his or her participation was valuable even without a CATI respondent. Participants without a CATI respondent were interviewed at the clinic using an expanded version of the In-Person Interview (IPI).

During the recruitment call, the Recruiter assessed whether travel arrangements were required and which clinic site would be the most accessible for the participant. In most cases, the clinic site selected was the one closest to the participant's usual residence. Seattle was selected as the clinic location for most participants living outside of the Pacific Northwest, for both participant convenience and typically lower airfare cost.

If a CATI respondent was named, the participant was called back after the CATI was completed to schedule the clinic visit. If no CATI respondent was named, attempts to schedule could begin immediately, depending on availability at the clinic location selected.

C.2.c. Confirmation of Agreement to Participate and Six Month Letter

In the Pilot Study, letters were sent to participants soon after they had agreed to participate, thanking them for agreeing to participate and explaining that they would be contacted to schedule a clinic appointment. During the Full Study, when the volume of letters to participants was extremely high and lag time between recruiting and scheduling could be six months or longer, the confirmation letter was replaced by a letter sent to all participants who had not yet been scheduled for a clinic within six months after the recruitment call. The purpose of the "six month letter" was two-fold. First, it served to assure participants that they would be contacted and that their participation was still very important to the study. Second, it enabled the study to be advised of address changes through the United States Postal Service change of address service.

C.2.d. Refusals and Second Attempts

While every reasonable effort was made to persuade each potential participant to participate, it was inevitable that some would refuse or be unable to participate. When a potential participant refused, the Recruiter asked them to complete a Refusal/Demographic Questionnaire (Appendix 5). Twelve demographic questions relating to race, ethnic origin, income, religion, and education level were asked to obtain a general profile of those who refused to participate or later withdrew from the study.

The Recruiter also completed a Refusal Assessment (Appendix 6) after the call to record the nature and strength of the refusal from the Recruiter's perspective. This information was used to determine if, and when, to re-contact the potential participant to have the best opportunity to convert the refusal to an agreement to participate. A second attempt to recruit was generally made unless the participant specifically requested that the HTDS not re-contact them, was hostile, or if it was clearly not possible for this person to participate in the study for reasons such as long term illness or disability. The default date for making second attempts was set at approximately three months after the first attempt.

C.2.e. Second Request for Participation

A second attempt letter, requesting participation and explaining the study, was sent within three to six months following the first recruitment contact, or after an appropriate amount of time based on information provided by the potential participant and the Recruiter. As with the initial contact, each letter also included a description of study participation and fact sheets. A Recruiter began attempts to call approximately 5-7 days after the second letter was sent. As with the first attempt call, a script was used as a guideline to ensure that all potential participants received the same information about the study. However, as with the initial attempt, the Recruiter's approach and responses were individualized to respond to the potential participant's questions and concerns.

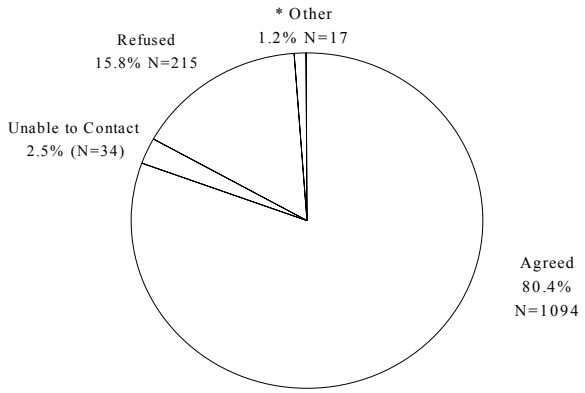
If the potential participant agreed to participate on the second attempt, the same steps were followed as for an agreement on the first attempt in requesting CATI respondent information and determining clinic location. If he or she refused, the Recruiter asked him or her to complete a Refusal Demographic Questionnaire (unless this had been completed during the first attempt call). The Recruiter also completed a Refusal Assessment after the call to record the reason for and strength of the refusal. No further recruitment attempts were made if a potential participant refused on the second attempt.

C.3. Outcome and Results

C.3.a. Results for the Pilot Study Sample

Figure V.C-1 summarizes the willingness of individuals in the Pilot Study Sample to agree to participate in the study. It should be noted that these numbers may vary slightly from those found in the Pilot Study Final Report of January 24, 1995, as some additional participants from the Pilot Study Sample were recruited after that time.

Figure V.C-1. Agreement to Participate for the Pilot Study Sample (N = 1360)



* Other: Unable to participate, Ineligible, Died prior to participation

Of the 1360 potential participants located alive, 1354 (99.6%) were sent letters requesting participation, and 1320 (97.1%) were contacted by telephone. One thousand ninety-four (80.4%) of those located alive agreed to participate (82.9% of those contacted by telephone). Eleven of those located alive were judged physically incapable of participating by a close relative or guardian, or were found to be otherwise unable to participate. Two hundred fifteen (15.8%) of the living located potential participants refused to participate in the study. Of those agreeing to participate, 49 (3.6%) refused on the initial attempt, but were re-contacted a second time and the refusal was converted to agreement. The participation rate remained remarkably constant over the course of the Pilot Study, fluctuating less than 3% in either direction over the last year of this phase of the study. Table V.C-1 summarizes the recruiting experience for the Pilot Study Sample.

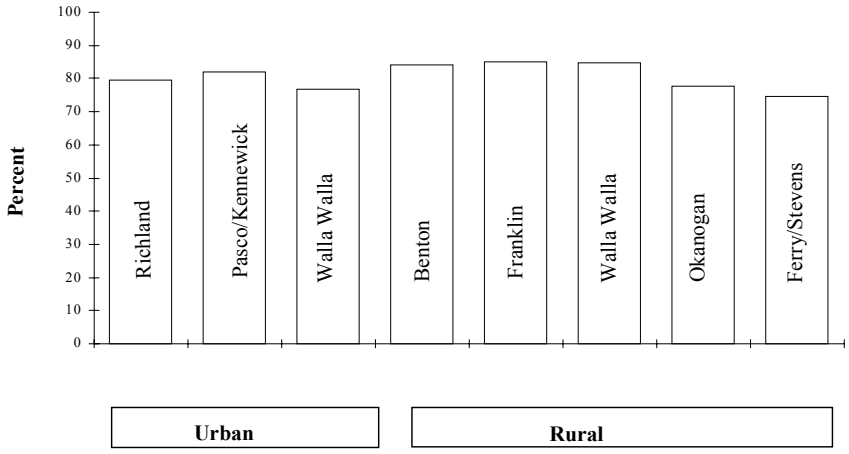
Table V.C-1. Summary of Agreement and Refusal for the Pilot Study Sample (N=1360)

Contact Status	No.	% of Contacted by Phone	% of Letter Sent	% of Total
Letter sent	1354	--	100.0	99.6
Unable to contact	34	--	2.5	2.5
Contacted by phone	1320	100.0	97.5	97.1
Agreed, final	1094	82.9	80.8	80.4
- on first attempt	1045	79.2	77.2	76.8
- on second attempt	49	3.7	3.6	3.6
Refused, final	215	16.3	15.9	15.8
Unable to participate	11	0.8	0.8	0.8
Died prior to participation	4	--	--	0.3
Ineligible	2	--	--	0.1
Located with no contact	0	--	--	0.0

In summary, the Pilot Study demonstrated that, once located and contacted by telephone, a large proportion of individuals would agree to participate in the study. Approximately 83% of those in the Pilot Study Sample who were contacted by telephone agreed to participate during the course of the study.

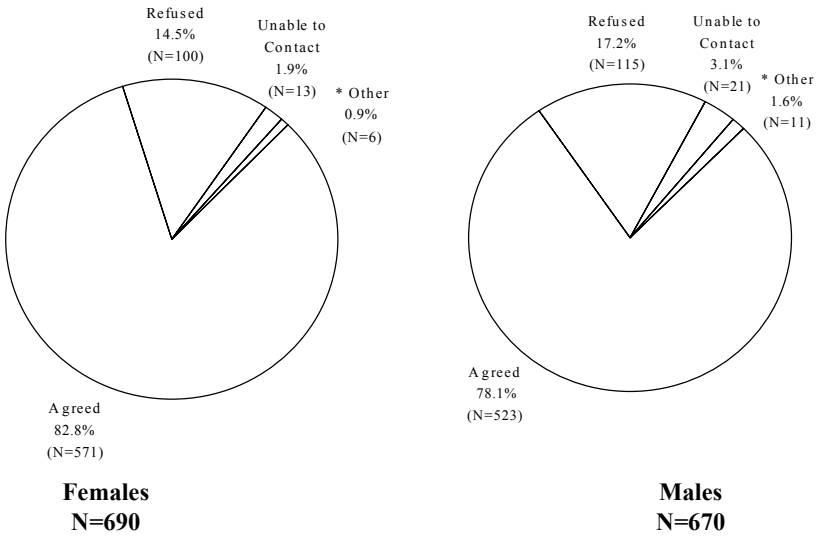
Figure V.C-2 shows the proportion of the Pilot Study Sample located alive who agreed to participate, by the eight geostrata used in the Pilot Study. The participation rate was uniformly high and relatively similar across the eight geostrata. The lowest percentage was among those born in Ferry and Stevens counties, located furthest from the Hanford Site (74.5%), and the highest was in Franklin County (84.9%), the area closest to the site.

Figure V.C-2. Agreement to Participate for the Pilot Study Sample, by Geostratum (N=1360)



Figures V.C-3 and V.C-4 show that willingness to participate also did not differ substantially according to sex (82.8% for females vs. 78.1% for males) or year of birth (78.8% to 82.4%).

Figure V.C-3. Agreement to Participate for the Pilot Study Sample, by Sex (N = 1360)



* Other: Unable to participate, Ineligible or Died prior to participation

Figure V.C-4. Agreement to Participate for the Pilot Study Sample, by Year of Birth (N = 1360)

Similarly, the area of current residence did not significantly influence the willingness to participate. The most common reasons for non-participation were "Not Interested" and "No Time," with 64.7% (139) of refusals and withdrawals falling into these two categories. The next most common reason for non-participation was illness or medical impairment/disability. Eighteen (8.4%) Pilot Study non-participants or their family member/guardian cited a medical condition, illness, disability, or impairment as the reason for not participating. Surprisingly, unwillingness to travel was only cited as the reason for non-participation by four potential participants outside the Northwest in the Pilot Study Sample. Other reasons given were opposition to the study, concern about the effect of participation on insurance coverage, advice from an attorney, and not having thyroid disease. Table V.C-2 shows the reasons for refusal or withdrawal from the study by geographic area of current residence for the Pilot Study Sample.

Table V.C-2. Reason for Refusal/Withdrawal for the Pilot Study Sample by Geographic Area of Current Residence (N=215)

Area of Current Residence	Reason for Refusal or Withdrawal														
	Not Interested or No Time		Illness or Impairment		Unwilling to Travel		Opposed to Study		Legal or Insurance Concerns		No Thyroid Disease		Other*		Total
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.
All WA	82	66.1	9	7.3	1	0.8	8	6.5	3	2.4	1	0.8	20	16.1	124
Seattle	12	75.0	1	6.2	0	--	0	--	0	--	0	--	3	18.8	16
Everett	1	33.3	0	--	0	--	1	33.3	0	--	0	--	1	33.3	3
Tacoma	3	60.0	0	--	0	--	0	--	0	--	0	--	2	40	5
Olympia	3	75.0	0	--	0	--	0	--	0	--	0	--	1	25.0	4
SW WA	5	83.3	0	--	0	--	1	16.7	0	--	0	--	0	--	6
Wenatchee	9	69.2	3	23.1	0	--	0	--	0	--	0	--	1	7.7	13
Yakima	2	25.0	1	12.5	1	12.5	0	--	1	12.5	0	--	3	37.5	8
Spokane	22	75.9	0	--	0	--	3	10.3	0	--	1	3.4	3	10.3	29
Tri-Cities	24	61.5	4	10.3	0	--	3	7.7	2	5.1	0	--	6	15.4	39
SE WA	1	100	0	--	0	--	0	--	0	--	0	--	0	--	1
Other NW	21	60.0	3	8.6	1	2.9	2	5.7	2	5.7	1	2.9	5	14.3	35
CA/HI	9	52.9	5	29.4	0	--	1	5.9	1	5.9	0	--	1	5.9	17
Southwest	11	91.7	0	--	0	--	0	--	0	--	0	--	1	8.3	12
Midwest	7	58.3	1	8.3	4	33.3	0	--	0	--	0	--	0	--	12
South	7	70.0	0	--	0	--	0	--	0	--	0	--	3	30.0	10
East	2	40.0	0	--	0	--	1	20.0	0	--	1	20.0	1	20.0	5
Other	0	--	0	--	0	--	0	--	0	--	0	--	0	--	0
Total	139	64.7	18	8.4	6	2.8	12	5.6	6	2.8	3	1.4	31	14.4	215

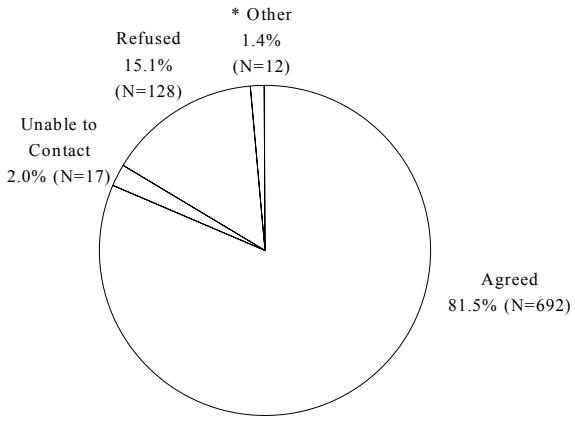
* Other: Includes family problems, personal reasons, distrust, lack of personal benefit, scheduling problems, refusal by a household member on the potential participant's behalf, and no reason given.

Overall, these results indicate that the methods developed for recruiting participants for the study were feasible and would result in relatively high levels of participation.

C.3.b. Results from the Transition Sample

Figure V.C-5 summarizes the willingness of individuals in the Transition Sample to agree to participate during the course of the study.

Figure V.C-5. Agreement to Participate for the Transition Sample (N = 849)



* Other: Unable to participate, Ineligible or Died prior to participation

Of the 849 potential participants located alive, 847 (99.8%) were sent letters requesting participation, and 831 (97.9%) were contacted by telephone. Six hundred ninety-two (81.5%) of those located alive agreed to participate. Eleven located potential participants were judged medically incapable of participating by a close relative or guardian contacted during the process, or were found to be otherwise unable to participate. One hundred twenty-eight of those located alive (15.1%) refused to participate in the study. Of those agreeing to participate, 33 (3.9%) potential participants refused on the initial attempt, but were re-contacted a second time and agreed to participate on the second recruiting attempt. Sixteen (1.9%) of those located to an address were unreachable by telephone and could not be recruited. Table V.C-3 summarizes the recruiting experience for the Transition Sample.

Table V.C-3. Summary of Agreement and Refusal for theTransition Sample (N=849)

Contact Status	No.	% of Contacted by Phone	% of Letter Sent	% of Total
Letter sent	847	--	100.0	99.8
Unable to contact	16	--	1.9	1.9
Contacted by phone	831	100.0	98.1	97.9
Agreed, final	692	83.3	81.7	81.5
- on first attempt	659	79.3	77.8	77.6
- on second attempt	33	4.0	3.9	3.9
Refused, final	128	15.4	15.1	15.1
Unable to participate	11	1.3	1.3	1.3
Died prior to participation	1	--	--	0.1
Ineligible	0	--	--	0.0
Located with no contact	1	--	--	0.1

In summary, the results in the Transition Sample did not differ appreciably from those in the Pilot Study. Of those individuals contacted by telephone, 83.3% agreed to participate in the study. A total of 2.0% (17) of those located to an address were not reachable by telephone.

Figure V.C-6 shows the proportion of Transition Sample potential participants located alive who agreed to participate, according to the five geostrata used in the Transition Sample. The participation rate was uniformly high and relatively similar across the five areas. The lowest percentage was among those born in Richland (77.6%), and the highest among those born in Pasco/Kennewick (86.7%).

Figure V.C-6. Agreement to Participate for the Transition Sample, by Geostratum (N=849)

Willingness to participate did not differ substantially according to sex (80.6% of females vs. 82.4% of males) or year of birth (78.7% to 86.3%) (Figures V.C-7 and V.C-8). Similarly, the area of current residence did not significantly influence the willingness to participate. Table V.C-4 summarizes the reasons for refusal by geographic area of current residence for the Transition Sample. Reasons for refusal did not vary substantially from the Pilot Study Sample experience.

Figure V.C-7. Agreement to Participate for the Transition Sample, by Sex (N=849)

Females
N=422

Males
N=427

* Other: Unable to particpate, Ineligible or Died prior to participation

Figure V.C-8. Agreement to Participate for the Transition Sample, by Year of Birth (N=849)

Table V.C-4. Reason for Refusal/Withdrawal for the Transition Sample, by Geographic Area of Current Residence (N=128)

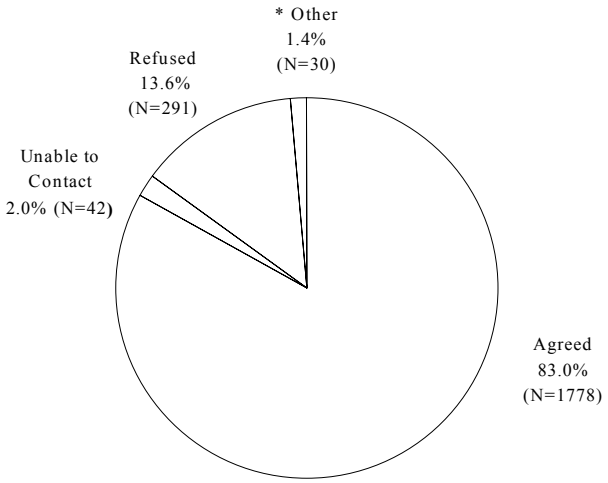
Area of Current Residence	Reason for Refusal or Withdrawal														
	Not Interested or No Time		Illness or Impairment		Unwilling to Travel		Opposed to Study		Legal Concerns		No Thyroid Disease		Other*		Total
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.
All WA	48	69.6	3	4.3	0	--	4	5.8	1	1.4	1	1.4	12	17.4	69
Seattle	6	60.0	0	--	0	--	0		0	--	0	--	4	40.0	10
Everett	2	100	0	--	0	--	0	--	0	--	0	--	0	--	2
Tacoma	2	100	0	--	0	--	0	--	0	--	0	--	0	--	2
Olympia	4	80.0	1	20.0	0	--	0	--	0	--	0	--	0	--	5
SW WA	1	33.3	0	--	0	--	1	33.3	0	--	0	--	1	33.3	3
Wenatchee	2	100	0	--	0	--	0	--	0	--	0	--	0	--	2
Yakima	4	57.1	1	14.3	0	--	1	14.3	1	14.3	0	--	0	--	7
Spokane	5	71.4	0	--	0	--	1	14.3	0	--	0	--	1	14.3	7
Tri-Cities	22	71.0	1	3.2	0	--	1	3.2	0	--	1	3.2	6	19.4	31
SE WA	0	--	0	--	0	--	0	--	0	--	0	--	0	--	0
Other NW	9	56.2	1	6.2	1	6.2	0	--	0	--	0	--	5	31.2	16
CA/HI	6	66.7	1	11.1	1	11.1	0	--	0	--	0	--	1	11.1	9
Southwest	2	50.0	0	--	1	25.0	1	25.0	0	--	0	--	0	--	4
Midwest	8	66.7	1	8.3	0	--	0	--	0	--	1	8.3	2	16.7	12
South	6	75.0	0	--	1	12.5	0	--	0	--	0	--	1	12.5	8
East	6	66.7	0	--	0	--	0	--	0	--	0	--	3	33.3	9
Other	0	--	0	--	0	--	0	--	0	--	0	--	1	100	1
Total	85	66.4	6	4.7	4	3.1	5	3.9	1	0.8	2	1.6	25	19.5	128

* Other: Includes family problems, personal reasons, distrust, lack of personal benefit, scheduling problems, refusal by a household member on the potential participant's behalf, and no reason given.

C.3.c. Results for the Full Study Sample

Figure V.C-9 summarizes the willingness of those located alive in the Full Study Sample to agree to participate during the course of the study.

Figure V.C-9. Agreement to Participate for the Full Study Sample (N = 2141)



* Other: Unable to participate, Ineligible or Died prior to participation

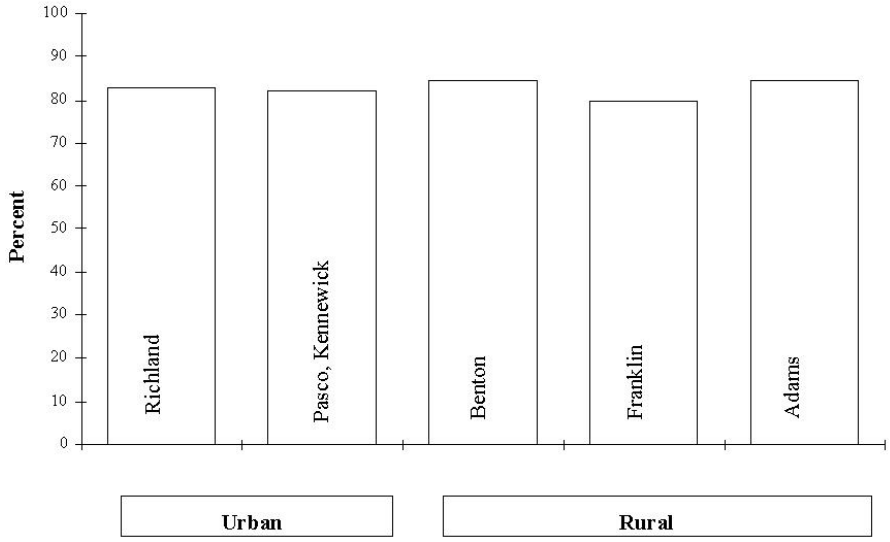
Of the 2141 Full Study Sample potential participants located alive, 2128 (99.4%) were sent letters requesting participation, and 2088 (97.5%) were contacted by telephone. A total of 1778 (83.0%) of those located alive agreed to participate. Nineteen (0.9%) of those located alive were judged medically incapable of participating by a close family member or guardian, or were found to be otherwise unable to participate. Two hundred ninety-one of those located alive (13.6%) refused to participate in the study. Of those agreeing to participate, 36 (1.7%) refused on the initial attempt, but were re-contacted a second time, and agreed to participate on the second recruiting attempt. Forty-two (2.0%) of those located to an address were unreachable by telephone. Table V.C-5 summarizes the recruiting experience for the Full Study Sample.

Table V.C-5. Summary of Agreement and Refusal for the Full Study Sample (N=2141)

Contact Status	No.	% of Contacted by Phone	% of Letter Sent	% of Total
Letter Sent	2128	--	100.0	99.4
Unable to contact	40	--	1.9	1.9
Contacted by phone	2088	100.0	98.1	97.5
Agreed, final	1778	85.2	83.6	83.0
- on first attempt	1742	83.4	81.9	81.4
- on second attempt	36	1.7	1.7	1.7
Refused, final	291	13.9	13.7	13.6
Unable to participate	19	0.9	0.9	0.9
Died prior to participation	11	--	--	0.5
Ineligible	0	--	--	0.0
Located with no contact	2	--	--	0.1

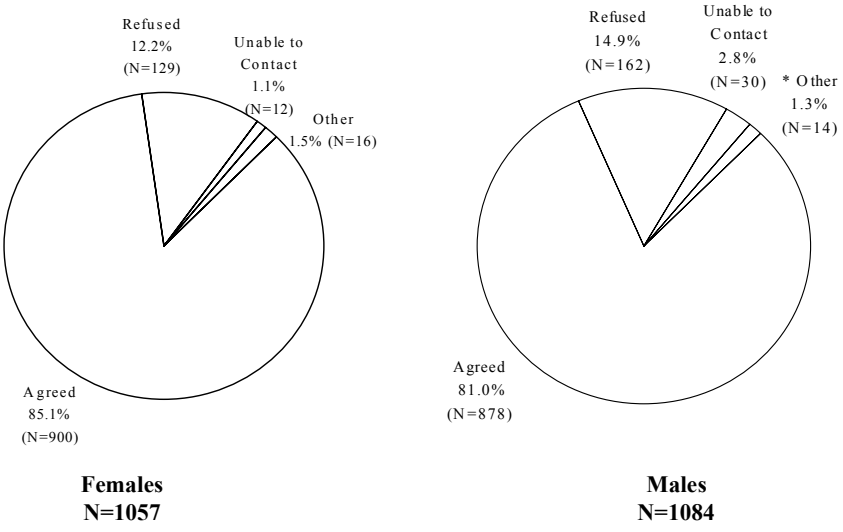
Figure V.C-10 shows the proportion of Full Study Sample members located alive who agreed to participate in the study, according to the five geostrata used in the Full Study Sample. The participation rate was uniformly high and relatively similar across the areas. The lowest percentage was among those born in Franklin County (80.0%), while the highest was among those born in Benton County (84.4%).

Figure V.C-10. Agreement to Participate for the Full Study Sample, by Geostratum (N=2141)



Willingness to participate also did not differ substantially according to sex (85.1% for females vs. 81.0% for males) or year of birth (77.5% to 89.6%) (Figures V.C-11 and V.C-12). Similarly, the area of current residence did not significantly influence the willingness to participate. Table V.C-6 shows the reasons for non-participation by geographic area of current residence. Reasons given were similar to those given in the Pilot and Transition Samples.

Figure V.C-11. Agreement to Participate for the Full Study Sample, by Sex (N = 2141)



* Other: Unable to participate, Ineligible or Died prior to participation

Figure V.C-12. Agreement to Participate for the Full Study Sample, By Year of Birth (N=2141)

Table V.C-6. Reason for Refusal/Withdrawal for the Full Study Sample, by Geographic Area of Current Residence (N=291)

Area of Current Residence	Reason for Refusal or Withdrawal														
	Not Interested or No Time		Illness or Impairment		Unwilling to Travel		Opposed to Study		Insurance Concerns		No Thyroid Disease		Other*		Total
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.
All WA	86	64.2	9	6.7	2	1.5	5	3.7	1	0.7	0	--	31	23.1	134
Seattle	10	55.6	1	5.6	0	--	1	5.6	1	5.6	0	--	5	27.8	18
Everett	4	57.1	1	14.3	0	--	0	--	0	--	0	--	2	28.6	7
Tacoma	6	75.0	0	--	0	--	0	--	0	--	0	--	2	25.0	8
Olympia	3	50.0	1	16.7	1	16.7	0	--	0	--	0	--	1	16.7	6
SW WA	6	66.7	0	--	0	--	1	11.1	0	--	0	--	2	22.2	9
Wenatchee	8	88.9	0	--	0	--	0	--	0	--	0	--	1	11.1	9
Yakima	8	72.7	1	9.1	0	--	0	--	0	--	0	--	2	18.2	11
Spokane	12	60.0	3	15.0	0	--	0	--	0	--	0	--	5	25.0	20
Tri-Cities	29	63.0	2	4.3	1	2.2	3	6.5	0	--	0	--	11	23.9	46
SE WA	0	--	0	--	0	--	0	--	0	--	0	--	0	--	0
Other NW	21	55.3	5	13.2	3	7.9	0	--	0	--	1	2.6	8	21.1	38
CA/HI	20	71.4	2	7.1	3	10.7	1	3.6	0	--	0	--	2	7.1	28
Southwest	17	63.0	5	18.5	1	3.7	2	7.4	0	0	0	0	2	7.4	27
Midwest	19	70.4	2	7.4	3	11.1	0	--	0	--	0	--	3	11.1	27
South	13	56.5	0	--	7	30.4	0	--	0	--	0	--	3	13.0	23
East	10	83.3	1	8.3	0	--	0	--	0	--	0	--	1	8.3	12
Other	1	50.0	0	--	1	50.0	0	--	0	--	0	--	0	--	2
Total	187	64.3	24	8.2	20	6.9	8	2.7	1	0.3	1	0.3	50	17.2	291

* Other: Includes family problems, personal reasons, distrust, lack of personal benefit, scheduling problems, refusal by a household member on the potential participant's behalf, and no reason given.

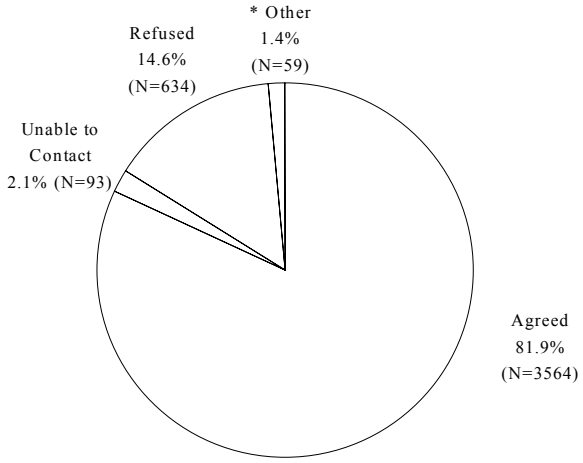
C.3.d. Results for the Entire Study

A summary of the agreement to participate for the entire study is provided in Table V.C-7 and shown in Figure V.C-13. In all, 4239 potential participants (97.4% of all living, located) were contacted by telephone to request participation. An additional 93 (2.1% of all living, located) were located to an address, and were sent one or more letters, but could not be contacted by telephone[2]. A total of 3564 potential participants (84.1% of those who were contacted by telephone, 81.9% of all located, living) agreed on either a first or a second attempt. Of those located alive, 634 (14.6%) refused to participate in the study.

Forty-one living located potential participants (0.9%) were determined to be unable to fully participate and were consequently not included in the study regardless of willingness to participate. Twenty-five were reported by others (parents, guardians or caregivers) to be incapable of participating due to mental or physical/medical disability. In these cases, contact with the person directly was not possible and could not be considered a refusal. Of the remaining 16, six were incarcerated out of state for the duration of the study; three were not opposed to participating, but were living outside of the U.S. and had no plans to return to the U.S. during the study; the remaining seven were either adopted, and/or did not have sufficient information regarding residence history of the birth mother or their early childhood to accurately assess residence/dose, and therefore would not have been evaluable (see section IV.B above for definition of evaluable participant).

[2]Either no phone number was available or multiple attempts to reach by phone resulted in no contact.

Figure V.C-13. Final Agreement to Participate for the Entire Study (N = 4350)



* Other: Unable to participate, Ineligible or Died prior to participation

Table V.C-7. Summary of Agreement or Refusal for the Entire Study (N=4350)

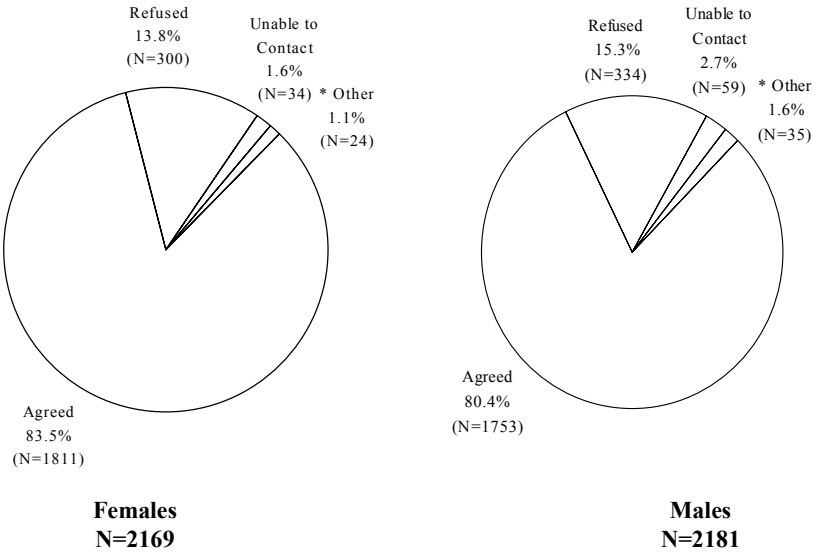
Contact Status	No.	% of Contacted by Phone	% of Letter Sent	% of Total
Letter Sent	4329	--	100.0	99.5
Unable to contact	90	--	2.1	2.1
Contacted by phone	4239	100.0	97.9	97.4
Agreed, final	3564	84.1	82.3	81.9
- on first attempt	3446	81.3	79.6	79.2
- on second attempt	118	2.8	2.7	2.7
Refused, final	634	15.0	14.6	14.6
Unable to participate	41	1.0	0.9	0.9
Died prior to participation	16	--	--	0.4
Ineligible	2	--	--	0.0
Located with no contact	3	--	--	0.1

Agreement to participate is shown by geostrata in Figure V.C-14. While those born in Ferry and Stevens Counties had the lowest agreement rate at 74.5%, willingness to participate did not otherwise differ substantially by geographic region of birth. Agreement rates from all other geographic strata ranged from 77.7-84.3%.

Figure V.C-14. Final Agreement to Participate for the Entire Study, by Geostratum (N=4350)

Agreement to participate is shown by sex for the entire study in Figure V.C-15. Women were slightly more likely to agree than men, 83.5% and 80.4% respectively.

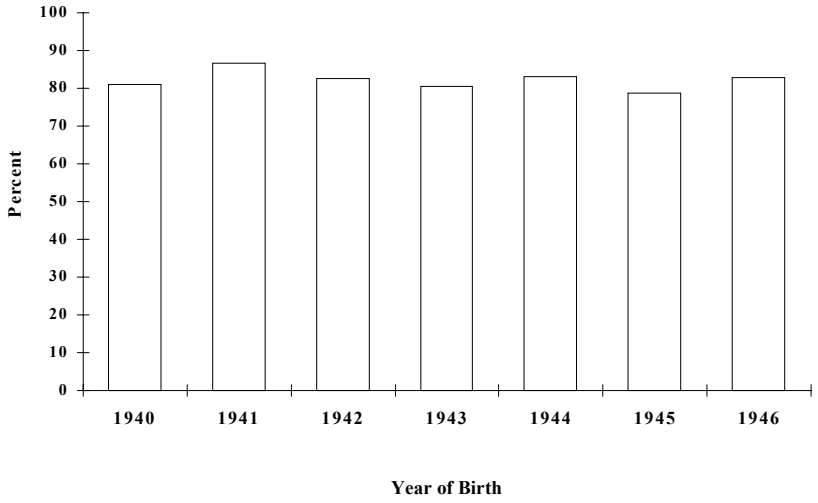
Figure V.C-15. Final Agreement to Participate for the Entire Study, by Sex (N = 4350)



* Other: Ineligible, unable, or died prior to participation

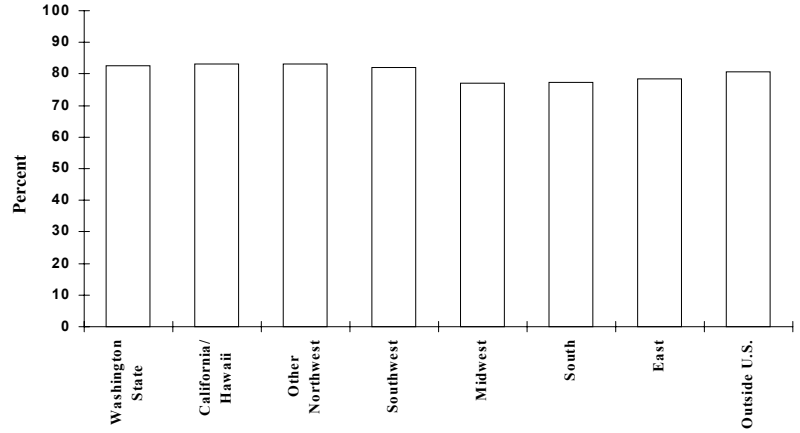
Willingness to participate did not vary appreciably by birth year, as shown in Figure V.C-16. Agreement rates range from 78.8% to 86.6% for the seven years of birth (1940-1946), with no apparent pattern.

Figure V.C-16. Final Agreement to Participate for the Entire Study, by Year of Birth (N=4350)



Agreement to participate by area of current residence is shown in Figure V.C-17. A slight variation was evident by region of the country. Agreement rates in the Midwest, Southern and Eastern portions of the US ranged from 77.0-78.4%, whereas in the western U.S. they ranged from 82.1-83.1%. Those living outside the U.S. had an agreement rate of 80.6%.

Figure V.C-17. Final Agreement to Participate for the Entire Study, by Geographic Region of Current Residence (N=4350)



The Regions in Figure V.C-17 were defined as follows:

Washington State	
Cal/Hawaii:	- California, Hawaii
Other Northwest:	- Alaska, Idaho, Montana, Oregon, and Wyoming
Southwest:	- Arizona, Colorado, Nevada, New Mexico, Texas, Utah
Midwest:	- Illinois, Indiana, Iowa, Kansas, Kentucky, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, Wisconsin
South:	- Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee
East:	- Connecticut, Delaware, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont, Virginia, Washington D.C., West Virginia

C.3.d.1. Agreement on First Attempt vs. Second Attempt (conversions)

Of those who "ever agreed" (3862), 95.9% (3704) agreed on the first attempt while 4.4% (172) agreed on the second attempt refusal conversion. Of all potential participants who "ever agreed," those agreeing on the second attempt were more likely to withdraw and/or never attend a clinic (39.5%), compared to those who agreed on the first attempt, but withdrew or never attended a clinic (9.6%). Nonetheless, it is still noteworthy that of the 172 potential participants whose initial refusal was converted to an agreement on a second attempt, 104 (60.5%) did eventually attend clinics, making up 3% of all participants attending clinics.

Table V.C-8 shows the reasons given for refusal or withdrawal, by geographic area of current residence. Overwhelmingly, "Not Interested" and "No Time" were the main reasons cited for non-participation, with 64.8% of all refusals falling into this category. The next highest category, at 7.6% (48 cohort members) was illness or impairment. In general, the reasons given did not vary significantly by area of current residence, although fewer Washington State residences cited unwillingness to travel as compared to those outside the state. Still, this reason accounted for only 30 (4.7%) refusals to participate.

Table V.C-8. Reasons for Refusal or Withdrawal for the Entire Study, by Geographic Area of Current Residence (N=634)

Area of Current Residence	Reason for Refusal or Withdrawal														
	Not Interested or No Time		Illness or Impairment		Unwilling to Travel		Opposed to Study		Legal or Insurance Concerns		No Thyroid Disease		Other*		Total
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.
All WA	216	66.1	21	6.4	3	0.9	17	5.2	5	1.5	2	0.6	63	19.3	327
Seattle	28	63.6	2	4.5	0	--	1	2.3	1	2.3	0	--	12	27.3	44
Everett	7	58.3	1	8.3	0	--	1	8.3	0	--	0	--	3	25.0	12
Tacoma	11	73.3	0	--	0	--	0	--	0	--	0	--	4	26.7	15
Olympia	10	66.7	2	13.3	1	6.7	0	--	0	--	0	--	2	13.3	15
SW WA	12	66.7	0	--	0	--	3	16.7	0	--	0	--	3	16.7	18
Wenatchee	19	79.2	3	12.5	0	--	0	--	0	--	0	--	2	8.3	24
Yakima	14	53.8	3	11.5	1	3.8	1	3.8	2	7.7	0	0	5	19.2	26
Spokane	39	69.6	3	5.4	0	0	4	7.1	0	--	1	1.8	9	16.1	56
Tri-Cities	75	64.7	7	6.0	1	0.9	7	6.0	2	1.7	1	0.9	23	19.8	116
SE WA	1	100	0	--	0	--	0	--	0	--	0	--	0	--	1
Other NW	51	57.3	9	10.1	5	5.6	2	2.2	2	2.2	2	2.2	18	20.2	89
CA/HI	35	64.8	8	14.8	4	7.4	2	3.7	1	1.9	0	--	4	7.4	54
Southwest	30	69.8	5	11.6	2	4.7	3	7.0	0	--	0	--	3	7.0	43
Midwest	34	66.7	4	7.8	7	13.7	0	--	0	--	1	2.0	5	9.8	51
South	26	63.4	0	--	8	19.5	0	--	0	--	0	--	7	17.1	41
East	18	69.2	1	3.8	0	--	1	3.8	0	--	1	3.8	5	19.2	26
Other	1	33.3	0	--	1	33.3	0	--	0	--	0	--	1	33.3	3
Total	411	64.8	48	7.6	30	4.7	25	3.9	8	1.3	6	0.9	106	16.7	634

* Other: Includes family problems, personal reasons, distrust, lack of personal benefit, scheduling problems, refusal by a household member on the potential participant's behalf, and no reason given.

C.3.d.2. Refusals

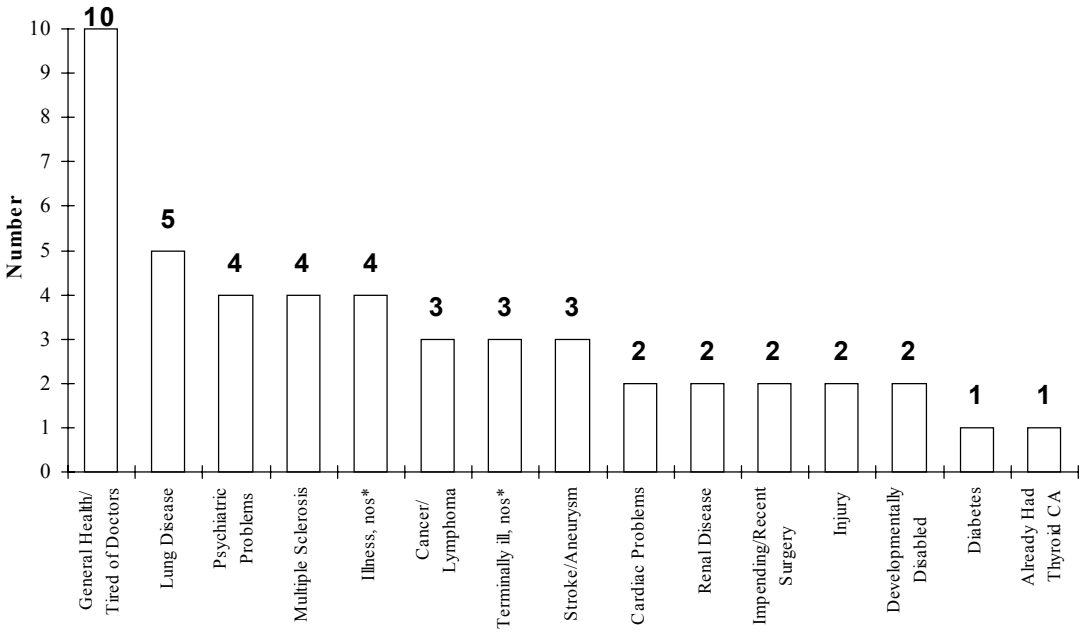
Among the final refusals (those either not recontacted or if recontacted, not converted to agreement), overwhelmingly, "not interested" and/or "no time" were the reasons given for most refusals, making up 31.9% and 33.0%, respectively. Other frequently cited reasons were "illness" (5.5%), "unwilling to travel" (4.7%), and "opposed to study" (3.9%). Particular efforts were made to accommodate potential participants who cited illness as a reason for being unable to participate, including repeated contacts, covering the cost of travel companion, special food, lodging and local travel accommodations. Nineteen (3.0%) responses indicated that a person other than the potential participant refused for the potential participant or discouraged them from participating. These were often spouses who were opposed to participation or who claimed to be responding on their spouse's behalf. While every effort was made to talk to the potential participant directly, the policy was not to pursue cases in which a family member would refuse for or not allow contact with the potential participant. In such cases, no further contact was attempted. This situation was thus considered a refusal to participate.

In eleven cases, the reason for non-participation was that the potential participant and Scheduler could not agree upon an acceptable clinic appointment. In most cases, this was simply due to the potential participant's extremely busy schedule at work and/or home, which precluded a genuinely interested participant from attending the clinic.

For those who refused or withdrew from the study due to illness, or were judged unable to participate due to impairment, the reason for their non-participation was recorded at the time of the refusal or withdrawal. The type of illness or impairment was recorded in the Recruiter's notes on the Refusal

Assessment Form. Figure V.C-18 summarizes the types of illnesses and impairments that precluded participation in the study. For those potential participants in the "Other" category reasons ranged from wanting the study to pay for their spouse to travel to the clinic with them for a vacation, to feeling that participation in a study with a politically controversial topic would conflict with their religious beliefs. In no case was current thyroid disease given as a reason for non-participation, however, one potential participant did state he/she did not wish to participate because he/she had already undergone a thyroidectomy for cancer and did not wish to have additional studies for this condition.

Figure V.C-18. Type of Illness/Impairment Precluding Participation for the Entire Study (N=48)



*nos = not otherwise specified

C.3.d.3. Success in Converting Refusals/Withdrawals by Reason for Refusal

With the exception of those who cited illness or gave no reason, there was relatively little variation in the success rate for converting initial refusals (shown in Table V.C-9, below). Success rates were higher for those who gave a reason of "no time" or "not interested" initially, and were highest for those who cited illness as the reason for initial refusal. The percent of conversions of those contacted a second time ranged from 0% for those reporting impairment as the reason for their refusal to 81% who refused for "other reasons".

Table V.C-9. Conversion to Agreement to Participate by Reason for Refusal for the Entire Study

		Recontacted		Converted to Agreement		
Reasons Refused	Total	No.	%	No.	% of Recontacted	% of Total
No time	250	147	58.8	41	27.9	16.4
Not interested	242	156	64.5	40	25.6	16.5
Illness	40	9	22.5	5	55.6	12.5
No reason given	34	9	26.5	4	44.4	11.8
Unwilling to travel	33	14	42.4	3	21.4	9.1
Opposed to study	29	15	51.7	4	26.7	13.8
Other person refused	22	12	54.6	3	25.0	13.6
Family problems	15	8	53.3	3	37.5	20.0
Impaired	13	3	23.1	0	0.0	0.0
Scheduling problems	11	0	0.0	0	0.0	0.0
No thyroid disease	6	2	33.3	0	0.0	0.0
Distrustful/suspicious	6	3	50.0	0	0.0	0.0
Advice of attorney	5	4	80.0	1	25.0	20.0
Insurance concerns	5	3	60.0	1	33.3	20.0
No personal benefit	5	0	0.0	0	0.0	0.0
Personal reasons	5	1	20.0	0	0.0	0.0
CATI upset respondent	3	2	66.7	0	0.0	0.0
Other reason	28	16	57.1	13	81.2	46.4
Total	752	404	53.7	118	29.2	15.7

* All who "ever agreed" on second attempt, including those who later withdrew

C.3.d.4. Success in Converting Refusals/Withdrawals by "Strength" of Refusal

It should be noted that whether or not a potential participant was re-contacted for a second attempt at recruitment was based on the Recruiter's or Participation Coordinator's discretion, potentially producing an inherently biased, "pre-selected" group of participants who were contacted for a second attempt. This may, in turn, affect the ability to accurately compare success in converting refusals by strength of refusal or other variables. With this in mind, it is still of interest to consider the success rates of refusal conversion by strength of initial refusal and by reason for refusal.

Fifty-six percent (245) of those classified as "firm" in their refusal on first attempt were re-contacted, while 60.9% (143) of those whose refusal was categorized as "mild" were contacted for a second attempt. Of those whose response was considered "hostile" on the first attempt, a total of 16 (20%) were designated for re-contact. These few potential participants were felt eligible for re-contact based on the point at which the refusal occurred (generally in the first seconds of the recruitment call).

Success in converting refusals to agreement to participate, when potential participants were re-contacted for a second attempt, did appear related to the strength of the initial refusal or withdrawal (classified by the Recruiter as mild, firm, or hostile). When contacted for a second attempt, 36.4% of participants whose refusal was categorized as "mild" agreed on a second attempt. In comparison, 24.5% of the participants who had been reported as "firm" in their refusal on the first attempt agreed on the second attempt. Interestingly, of those judged "hostile" on initial contact who were re-contacted, 37.5% agreed on the second attempt. This was, however, a small and highly select group, which would not represent "hostile" refusers as a whole.

The variation in success rates for conversion of refusals when comparing all potential participants (whether re-contacted or not) by strength of refusal was similar between the "mild" and "firm" refusals, 22.1% and 13.7%, respectively. Table V.C-10 shows conversion to agreement by strength of refusal.

Table V.C-10. Conversion to Agreement to Participate by Strength of Refusal for the Entire Study

Strength of Refusal	No.	Re-contacted	%	Converted	% Converted of Re-contacted	% Converted of Total Refused
Mild	235	143	60.9	52	36.4	22.1
Firm	437	245	56.1	60	24.5	13.7
Hostile	80	16	20.0	6	37.5	7.5
Total	752	404	53.7	118	29.2	15.7

While the attempts to convert initial refusals or withdrawals appear to be more successful when re-contacting those whose refusals were classified as mild, with 13.7% of all "firm" refusals/withdrawals agreeing to participate on the second attempt, it also seemed worthwhile to attempt to convert most potential participants regardless of perceived strength of first refusal (with the exception of truly hostile potential participants who were generally not re-contacted).

C.3.e. Conclusions

Efforts to recruit participants for this study were successful and met or exceeded initial expectations. Although participation required a substantial effort on the part of the participant and his/her family, these rates of success indicate a substantial degree of general willingness of those selected to participate in this study.

An important part of this success can be attributed to highly trained recruiting staff who was able to respond appropriately to potential participants' concerns. The approach of sending detailed study information prior to contacting potential participants seemed to work well for the purposes of recruiting. In addition, re-contacting many potential participants who refused on the first attempt (or withdrew), resulted in substantial refusal conversions and a 3% increase in those ultimately attending a clinic.

D. Computer Assisted Telephone Interview

D.1. Background

Two basic approaches were considered for collecting information about study participants' early years of life: 1) a personal interview with one respondent and other members of the family present; and 2) a telephone interview with a respondent and other family members connected by a conference call. The approach of conducting a personal interview was deemed not to be feasible due to the logistical complexities of organizing such interviews all over the country and the very high costs that would be associated with such a process. A decision was made to proceed with the development of a Computer Assisted Telephone Interview (CATI). The idea of involving more than one person was later incorporated, to a limited degree, when special circumstances dictated that an additional person (or persons) would enhance the recall of specific information (e.g., cow feeding patterns).

D.1.a. Objectives of the Interview

The primary objective of the CATI was to collect information that would be used as input for calculating a radiation dose to the thyroid gland from Hanford's ^{131}I for each study participant, as well as information about other radiation exposures and diagnoses of thyroid disease experienced by the participant. Secondary objectives of the CATI component of the study were to: 1) interview a person knowledgeable about each participant's early life (e.g., someone who could answer questions about the whereabouts, circumstances, and habits that an individual could not be expected to know about his/her very early years); and 2) assure that the accuracy and integrity of the data collected were of acceptable quality.

D.1.b. Historical Perspective and Special Challenges

Prior to the time that the HTDS began developing a CATI for dose determination purposes, the CATI technique had been used extensively for several years by many organizations, primarily to conduct telephone surveys for health and opinion research. A CATI is conducted by an interviewer who reads the survey text and questions from a computer screen. As the respondent provides each answer, the interviewer enters the response into the computer, and the response immediately becomes part of the permanent database. The computer program is designed to show the next question on the screen that should be asked, based on the previous answer(s).

Surveys conducted with CATI are generally quite straightforward, and they are usually structured in such a way that the questions are formatted for multiple choice, true/false - agree/disagree, and short answer responses. The total interview time is seldom greater than twenty minutes, and the respondent does not prepare in advance for the interview. CATI is often used in conjunction with "cold calling" to identify respondents willing to spend a few minutes on the telephone participating in a survey.

In the initial stages of planning the HTDS CATI, it became apparent that the interview would be far more complex than is typical of the CATI format. It appeared that the "state of the art" for CATI methodology did not incorporate many of the key features that would be required for the HTDS CATI. Several characteristics of the HTDS posed special challenges to developing a workable CATI.

First, much of the information required pertains to events that happened between forty and fifty years before the interview. Furthermore, much of the information in the interview could be considered rather mundane in that the questions would need to refer to events and circumstances of daily life. Clearly, asking people to recall such detailed information from so long ago would present very special challenges.

Second, a large volume of information would need to be collected during the interview. It would be necessary to develop a structure that would organize the various types of data collected, while

accommodating a wide range of life circumstances among the study participants. For example, some participants were born in the study area while their parents resided there only temporarily, perhaps for only a few weeks, before moving out of the Northwest. Such an interview would yield a relatively small amount of data. Other participants were born and lived their entire lives in the area, perhaps at multiple residences. An interview about such an individual could produce a much larger amount of data. It would be important to have computer software that could adequately adjust to the very different circumstances that would likely arise, and the variations in the amount of data collected. Third, and related, it would be critically important that the system be capable of managing complex skip patterns, and allow for on-line consistency checks and the ability to correct entries on-line.

D.2. Content and Design of the CATI

The plan for CATI described in the HTDS Protocol provided the rationale for the content of the interview and identified its components. It was designed to collect information from the early years of the participants' lives, including time *in utero,* from 1944 to 1957. The period of greatest interest, with regard to exposure to radioactive iodine, was each participant's early childhood. The interview was "location-driven" so that the information collected was specific to locations and periods of time directly relevant to the radiation releases from Hanford.

The following topic areas were included in the CATI: 1) a residential history of the participant from birth through 1957, and for the mother while pregnant with and breastfeeding the participant; 2) sources of milk consumed by the participant from birth through 1957, and for the participant's mother while pregnant and breastfeeding (including commercial milk producers and private sources, for both cow's and goat's milk; 3) milk consumption patterns for the participant from birth through 1957, and for the mother during pregnancy and breastfeeding; 4) other patterns of food consumption, including green and leafy vegetables, fresh fruit and free-range chicken eggs, for the participant from birth through 1957, and for the mother while pregnant and breastfeeding. In addition, medical history information was obtained for both the mother and the participant, including the following: 1) thyroid diseases and selected other medical conditions diagnosed and treated in the participant; 2) history of radiation exposures, either diagnostic or therapeutic, for the participant, and for the mother during pregnancy and breastfeeding. The name of the treating physician for these conditions and treatments was obtained when possible.

The CATI was developed in cooperation with a number of individuals and groups, including the Technical Steering Panel of the Hanford Environmental Dose Reconstruction Project and Battelle Pacific Northwest Laboratories, staff at the CDC, scientists who had conducted similar studies (e.g., Dr. Lynn Lyon at the University of Utah), and experts in survey and cognitive research.

After extensive investigation of available software options, the INGRES software package was selected as the basis for developing the CATI. INGRES provided a relational database structure, which was judged to be essential for the type of system envisioned, and contained many of the technical features needed to accommodate a complex interview with on-line quality control.

The CATI was administered to the participant's mother, or other person knowledgeable about the participant's early years, by specially trained interviewers. The interview was recorded on audiotape with the respondent's permission, so that a permanent record, independent of the computer system, would be created. The recording could be used for back-up to the computer system, for training and quality control monitoring of interviewers, and for clarification of information provided during the interview.

D.2.a. Development of a Cognitive Approach to Enhance Long-term Recall

From the initial stage of questionnaire development it was apparent that making the interview successful would depend largely on the ability of respondents to accurately report detailed information

about their child (or sibling) from very long ago. In July 1990 a workshop was held to consider how the questionnaire and the process of conducting the CATI interview could be modified to include as many characteristics of a cognitive interview as possible. Participants in the workshop included Dr. Donald Dillman, a sociologist and leading national authority on survey research from the Washington State University at Pullman; Dr. Ronald Fisher, a cognitive psychologist from Florida International University; and Dr. David Price, an agricultural economist and member of the Hanford Environmental Dose Reconstruction Project Technical Steering Panel from Washington State University, the four HTDS investigators, and key HTDS staff (Project Manager, Programmer, Field Operations Supervisor). Dr. Dillman's expertise in interview data collection has been utilized by the U.S. Department of the Census, while Dr. Fisher's work has been used extensively to assist in both criminal investigation and investigations of food-borne illness.

The cognitive interview is a technique developed to enhance recall. It is based on principles of cognition and memory retrieval theory. In the cognitive interview, it is important to mentally take the respondent back to the time period in question, and have them remember as much about that time as possible. As more memories of the time in question are recalled by the respondent, the likelihood of remembering answers to specific questions increases. Thus, for an interview regarding food consumption patterns such as the HTDS CATI, one would want to guide the respondent to remember not only major events or favorite songs of the time, but what the kitchen where the food was prepared looked like, and where food was purchased. These principals of the cognitive interview, with extensive preparation by the respondent, differ greatly from the standard epidemiologic interview. In most epidemiologic studies, great care is taken to ensure that the respondent does not prepare in advance to answer questions. Such preparation, it is felt, could produce bias in that those who are ill may be more likely to prepare and report exposure than those who are not. However, these studies also generally do not ask such specific questions about daily life events so many years after the fact.

D.2.b. Development and Testing of the CATI

During the spring of 1991 the first field testing of a paper version of the questionnaire took place in the Tri-Cities area. Individuals who had offered to help the study in some way were asked to participate in the testing. Three interviews were conducted with people in their homes. These individuals closely fit the profile of a CATI respondent. Generally, they were in the same age range as the parents of study participants, and they had children who were born during nearly the same years as study participants. Care was taken not to include individuals who could possibly be asked later to participate in the actual study.

Major conclusions drawn from this field test included the following: 1) it was too difficult for respondents to look at maps and determine the exact locations of residences; 2) the memory prompts previously developed were helpful, but needed to be expanded to encourage advance preparation by the respondent; 3) asking respondents to identify all residences during the interview without preparing beforehand was too difficult; and 4) giving the respondent the opportunity to prepare ahead of time for the interview was very important, and would be a major determinant in obtaining a successful interview.

During the early summer of 1992, Dr. John Tarnai, a sociologist from Washington State University and colleague of Dr. Don Dillman, began working with the HTDS staff on expanding and refining the memory materials that would be provided to respondents in preparation for the interview. As a result of the field testing concluded in 1991, it was decided to ask respondents to provide a written residence history to be mailed to the study office prior to conducting the interview. One goal of the memory materials was to encourage recall for completion of the residence history by providing information about events that happened during each year of interest. World, national, and local events, as well as popular songs, movies, and trends from each year were included to help provide a frame of reference that would help direct memory to many years ago.

Additional memory materials were developed to help the respondent prepare for answering the interview questions. Background information was provided to encourage memory about specific topics. For example, the dates of VE Day and the death of President Roosevelt were provided as general reference dates, while the beginning of war rationing and the Tri-Cites Memorial Day flood of 1948 were added to focus on local events which might have impacted food consumption practices. The memory materials were organized into a booklet that was to be sent with the residence history questionnaire. In addition, the text of the interview was refined to include references to specific parts of the memory materials at key points during the interview.

A second and more extensive field test was conducted during July and August of 1992. Telephone interviews were conducted with parents, friends, relatives of HTDS and other FHCRC staff members, and a few individuals recruited from local senior citizens centers. All respondents were similar in age to the parents of study participants. Precautions were taken to ensure that none of the individuals involved could later be asked to participate in the study. Fifteen individuals participated as respondents in this effort.

This round of field testing consisted of two parts. Interviews were completed with about half the individuals, and they were then asked to provide feedback about the interview experience. The primary finding was that the volume of materials provided for memory recall purposes was overwhelming. As a result, the materials were divided into two parts. The first booklet, titled the Calendar of Events (Appendix 7), would accompany the Residence History Questionnaire (Appendix 8) that respondents would complete and mail back prior to the interview. The second, titled the Interview Booklet (Appendix 9), was designed to contain information that would help prepare for answering the interview questions. The Interview Booklet was to be mailed a few days after the Calendar of Events and Residence History Questionnaire were sent.

The revised materials, sent out in two separate mailings, were used during the second part of field testing. These later interviews confirmed that dividing the materials was easier for the respondents, as the volume of information was not so intimidating. In response to comments from the second group that it was difficult to foresee what the questions in the interview would be like, a sheet of sample questions was developed. An additional page of materials entitled “Meet the Johnsons,” presented a profile of a typical family, then gave examples of questions from the interview with the appropriate responses, based on information provided in the profile. This sheet was enclosed with the Interview Booklet, and is included here as Appendix 10.

Additional smaller refinements to the questionnaire text were made during the early fall of 1992, as a result of the CATI training and practice interviews (Appendix 11).

D.2.c. Final Process and Procedures

D.2.c.1. Conducting the Interview

Each participant recruited for the study was asked to identify a respondent for the CATI as described in section V.C.2.b above. This person was to be knowledgeable regarding the participant’s early life and eating habits, able to perform the required preparation for the interview, and able to respond to the questions over the phone during a conversation that could be over an hour in length.

Once the respondent was identified, a letter was sent informing her or him that the participant had asked that they complete this portion of the study. This letter was followed by a phone call from the Interviewer to explain the process and obtain consent to do the CATI. If the respondent declined to do the interview, the participant was recontacted to determine if another respondent was available. If the respondent consented to the interview, the Residence History Questionnaire and Calendar of Events were sent to them to complete. The Residence History Questionnaire was to be sent back to the Interviewer, and once received, the Interviewer called the respondent to review the information and schedule the actual

interview. With the respondent's consent, interviews were recorded on audiotape for quality control and interviewer training purposes.

Once the interview was completed, the Interviewer updated the tracking system, so the participant could be scheduled for a clinic appointment. A thank-you letter was sent to each respondent in appreciation of his or her participation in the study.

D.2.c.2. Quality Control

Quality control for the CATI was first addressed in the thorough training given to each interviewer prior to performing actual interviews. Each interviewer was provided extensive training on both the interview instrument as well as the computer system required to administer the CATI. Over the course of the study, seven interviewers were trained to conduct the CATI. Each interviewer received written materials including a flow diagram of the entire interview, a question-by-question training manual, and a manual covering interviewing techniques such as appropriate probing and responses to respondent questions. In addition, they received documentation of the CATI program, special training on making data corrections during the interview (when respondents changed their minds regarding a previous answer), and a procedure manual outlining the CATI process from initial contact to completion of the interview.

When the study first began, the original three interviewers traveled to Washington State University in Pullman for training in the cognitive interview technique. Two additional interviewers underwent this training later in the study, while the final two interviewers hired received this portion of the training from experienced HTDS interviewers.

Interviewers continued their training by conducting the interview with HTDS staff, family and friends. This was followed by practice interviews with volunteers (often the parents of HTDS staff members) who had children in the age range of study participants. Tapes of these "practice" interviews were reviewed with the CATI Supervisor and experienced interviewers for feedback on technique and accuracy. Later in the study, new interviewers began their training by listening to previous interviews with an experienced interviewer.

Throughout the study, the CATI supervisor listened to tapes of the interviews as part of the quality control plan. Checks of the data entered during the interview were compared to the answers given on the tape. Any necessary data corrections were performed by the systems analyst/programmer. Feedback on any errors found was given to individual interviewers by the CATI Supervisor. In addition, early in the study, tapes were copied and forwarded to Dr. John Tarnai and Ms. Ellen Lammiman of Washington State University, Pullman. These recordings were reviewed for interviewer technique in assisting recall of participants, appropriate probing questions, and consistency. Feedback from Dr. Tarnai and Ms. Lammiman was forwarded to the interviewers as part of the ongoing assessment of their work. Sampling of tapes was performed on a random basis for two interviews per week during the first six months of the study. Additional tapes were monitored following the training of new interviewers, or when specific issues were found. Random checks continued throughout the study at a rate of approximately one interview per week.

Quality assessment of the respondent's ability to answer the interview questions was also performed. Following each section of the interview, interviewers recorded their assessment of how reliable the responses were for those questions using the categories of High, Generally Reliable, Questionable, or Unreliable. These assessments were based on whether the respondent seemed fairly certain of the responses, appeared to be guessing or asking the interviewer for help in making the "correct" response, and whether the responses were consistent or contradictory. At the end of the interview, the interviewer also recorded her or his overall assessment of the reliability of the responses, and of the respondent's level of cooperation (Very Good, Good Fair, or Poor).

D.3. Outcome and Results

D.3.a. Pilot Study Results

CATIs were completed for 797 (85.1%) of the 937 participants in the Pilot Study Sample who identified a CATI respondent. Of the 1063 Pilot Study participants who completed the clinic, 756 (71.1%) had a complete CATI interview. Forty-one participants withdrew from participation after the CATI was completed. In 14 instances, CATI Interviewers deemed the quality of the data provided by respondents too poor to be considered reliable. Expanded interviews were performed at the clinic for these participants.

D.3.b. Transition Sample Results

Of the 536 participants in the Transition Sample who identified a CATI respondent, interviews were completed for 458 (85.4%). Of the 664 Transition Sample participants who completed the clinic, 429 (64.6%) had a CATI. Twenty-nine participants withdrew from the study after a CATI was completed. In two instances, CATI Interviewers determined that the quality of the data provided by respondents was too poor to be considered reliable. Expanded interviews were performed at the clinic for these participants.

D.3.c. Results for the Full Study Sample

CATIs were completed for 1011 (81.6%) of the 1239 participants in the Full Study Sample who identified a CATI respondent. Of the 1720 Full Study Sample participants who completed the clinic, 948 (55.1%) had a CATI. Sixty-three participants withdrew from the study after a CATI was completed. In 13 instances, CATI Interviewers determined that the quality of the data provided by respondents was too poor to be considered reliable. Expanded interviews were performed at the clinic for these participants.

D.3.d. Overall Results for the Entire Study

Of the 2712 participants who identified a CATI respondent in the entire study, interviews were completed for 2266 (83.6%). Of the 3447 eligible participants who completed the clinic, 2133 (61.9%) had a CATI. One hundred-thirty-three participants withdrew from the study after a CATI was completed. In 29 of the 2133 instances, CATI Interviewers determined the quality of the data provided by respondents was too poor to be considered reliable. Expanded interviews were performed at the clinic for these participants.

D.3.e. Conclusions

The percentage of CATIs completed for participants declined with each successive phase of the study. This can probably be attributed to the fact that respondents were somewhat older as the study progressed, especially following the Full Study Sample. Participants born in 1940 and 1941 were included at this time, and this small difference in birth years may have contributed to the decrease in the overall percentage of CATIs completed. Table V.D-1 shows the number of CATIs completed by year of participant's birth.

Table V.D-1. Final Outcome of CATI by Participant's Year of Birth for the Entire Study (N =2712*)

Year of Birth	Respondent Identified	CATI Completed* No.	CATI Completed* % of those w/Respondent Identified	CATI Completed for those Attending a Clinic	% of those Attending a Clinic with CATI
1940	164	135	82.3	128	52.0
1941	170	136	80.0	129	45.3
1942	366	317	86.6	299	63.2
1943	438	348	79.5	326	58.2
1944	742	608	81.9	569	62.7
1945	504	436	86.5	412	67.4
1946	328	286	87.2	270	74.0

* Includes all CATIs completed, whether acceptable for dose determination or not.

D.3.e.1. Quality of the Data

Overall data quality was very high as reported by the interviewers. Tables V.D-2 through V.D-4 show the assessment of data reliability as reported by the CATI interviewers. Responses were judged to be of high quality or generally reliable for most interviews for most sections. Responses in the section related to the participant's milk and dietary consumption history were judged by the interviewers to be questionable in approximately 9% of the interviews. Not surprisingly, the main reason cited for questionable or unreliable responses was unclear memory. The interviewer judged the respondent's cooperation to be good or very good in over 94% of the interviews.

Table V.D-2. Interviewer's Overall Assessment of Reliability of Responses to CATI (CATIs Used for Dose Estimation Only) for the Entire Study (N=2123)

	Overall		Milk Source Data		Mother's Milk Consumption and Dietary Data		Participant's Milk Consumption and Dietary Data		Mother's Medical History		Participant's Medical History	
Response Quality	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
High	410	19.3	387	18.2	377	17.8	297	14.0	1178	55.5	951	44.8
Generally reliable	1523	71.7	1570	74.0	1552	73.1	1483	69.9	910	42.9	1118	52.7
Questionable	170	8.0	39	1.8	53	2.5	193	9.1	23	1.1	44	2.1
Unreliable	11	0.5	0	0	1	0	10	0.5	3	0.1	4	0.2
Unknown	9	0.4	127	6.0	140	6.6	140	6.6	9	0.4	6	0.3

Table V.D-3. Main Reasons for Unreliable or Questionable Responses to CATI (CATIs Used for Dose Estimation Only) for the Entire Study (N=2123)

	Overall		Milk Source Data		Mother's Milk Consumption		Participant's Milk Consumption and Dietary Data		Mother's Medical History		Participant's Medical History	
Response Quality	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Unclear memory of events	136	6.4	34	1.6	41	1.9	141	6.6	21	1.0	35	1.6
Uncertain understanding of questions	9	0.4	1	0	2	0.1	17	0.8	0	0	1	0
Hurried responses	8	0.4	1	0	2	0.1	11	0.5	0	0	1	0
Other	28	1.3	3	0.1	8	0.4	33	1.6	5	0.2	9	0.4
Don't know	0	0	0	0	1	0	1	0	0	0	2	0.1
Not applicable*	1942	91.5	2084	98.2	2069	97.5	1920	90.4	2097	98.8	2075	97.7

* Response quality High, Generally Reliable or Unknown

Table V.D-4. CATI Interviewer's Assessment of Respondent's Cooperation (CATIs Used for Dose Estimation Only) for the Entire Study (N = 2123)

Respondent's Cooperation	No.	%
Very good	1511	71.2
Good	496	23.4
Fair	99	4.7
Poor	8	0.4
Not answered	9	0.4

It was anticipated from the beginning of the study that participants' mothers would be the most reliable respondents for the majority of the interview questions, as mothers would be most familiar with the participant's dietary habits and medical histories. This was generally the case. Table V.D-5 shows the relationship of the respondent to the study participant, while Table V.D-6 shows the quality of the CATI data by the respondent's relationship to the participant.

Table V.D-5. Relationship of CATI Respondent to Participant for the Entire Study

	All Persons Who Agreed To Participate		Living Evaluable Participants			
	All Interviews (N=2268)		All Interviews (N=2133)		Interviews Used as Source of Dosimetry Data (N=2123)	
Relationship to Respondent	No.	%	No.	%	No.	%
Birth mother	1674	73.8	1577	73.9	1568	73.9
Adopted mother	8	0.4	6	0.3	6	0.3
Father	167	7.4	158	7.4	158	7.4
Sister	289	12.7	270	12.7	270	12.7
Brother	89	3.9	82	3.8	81	3.8
Aunt	29	1.3	28	1.3	28	1.3
Uncle	4	0.2	4	0.2	4	0.2
Other relative	3	0.1	3	0.1	3	0.1
Family friend	5	0.2	5	0.2	5	0.2

Table V.D-6. Quality of CATI Data by Respondent's Relationship to Participant for the Entire Study (N =2268)

Overall Response	Birth Mother		Adopted Mother		Father		Older Sister		Older Brother		Other Family Member		Family Friend	
Quality	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
High	328	19.6	0	0	28	16.8	59	20.4	23	25.8	1	2.8	1	20.0
Generally reliable	1203	71.9	7	87.5	121	72.5	208	72.0	54	60.7	23	63.9	1	20.0
Questionable	125	7.5	1	12.5	14	8.4	18	6.2	12	13.5	12	33.3	3	60.0
Unreliable	7	0.4	0	0.0	3	1.8	4	1.4	0	0.0	0	0.0	0	0.0
Unknown	11	0.7	0	0.0	1	0.6	0	0.0	0	0.0	0	0.0	0	0.0
Total	1674		8		167		289		89		36		5	

D.4. Attempts to Administer the CATI to Respondents for Deceased Potential Participants

D.4.a. Development of a Revised CATI for Deceased Potential Participants

The HTDS Protocol stated that CATIs would be conducted for deceased potential participants, using the CATI respondent as a surrogate for the potential participants in collecting information contained in the In-Person Interview as well. A separate CATI instrument was developed for this purpose in the late summer of 1994. The questionnaire was an expansion of the original CATI, adding questions that were part of the In-Person Interview administered to living participants. These included questions about occupational history, smoking history, and demographics. It was recognized that, depending on the age of the potential participant at death, some questions would not be pertinent. Only those germane to the participant's life circumstances would be asked during the interview.

Special memory materials and interview preparation materials were developed for use with the revised CATI for deceased potential participants. Although similar to those for the interviews conducted for living participants, there were some differences in content: 1) the residence history information was collected from birth to death instead of through 1957 only; 2) a cause of death information questionnaire was added for the respondent to complete; and 3) sections about occupational history and smoking history were added to the Interview Booklet.

D.4.b. Conducting a CATI for Deceased Potential Participants

Thirty-three potential participants known to be deceased were selected from HTDS tracing records during the fall of 1994 to test the revised CATI process and instruments. Those selected included a large number of potential participants who died in infancy. The objective was to select cases for whom the interview would be comparatively uncomplicated. Letters of approach were sent to the respondents, and recruiting of the respondents was begun after about one week, as with living study participants. The CATI Interviewers found that the respondents had difficulties discussing the deceased potential participants and their lives.

The overall refusal rate among respondents for deceased potential participants was about 55% (18 of 33) at the end of this short pilot project. Although some respondents initially agreed to participate, as the process unfolded they found they could not proceed. They reported that the experience was just too painful for them to continue. Interviews were eventually completed for 15 of the 33 (45%) deceased potential participants.

Based on this pilot experience, it was decided that pursuing such an approach would be difficult for respondents and staff, and would not be likely to produce data of sufficient quality to be useful in estimating doses for deceased potential participants. Thus, in March of 1995, the decision was made not to attempt a CATI interview for deceased potential participants. The reasons for this decision were presented to the CDC and the HTDS Advisory Committee, who agreed that further attempts to perform CATIs for deceased potential participants were not warranted.

D.5. Success of the CATI Component

Despite significant obstacles, the CATI component of the study was quite successful, not only in terms of completion of interviews, but in the success of the programming and logistical aspects of the CATI. Because no existing CATI software was available which would accommodate the needs of the dosimetry system, it was necessary to identify software that would be suitable for creating a custom interview to satisfy the requirements of the HTDS. This task was undertaken by Mr. Mark Saporito, Systems Analyst and Programmer for the HTDS, using the INGRES relational database program. While developing such a program required extensive lead-time and testing, it also allowed for a system which could be completely matched to the needs of the study, both in terms of the type of information gathered and the use of the cognitive interview.

The idea of using a cognitive approach added significantly to the development time of the CATI as well. Because the data being sought were from such a distant time period and revolved around fairly mundane activities of daily living, the cognitive approach was extremely important in eliciting accurate information from respondents. There was, however, an equally important need to refrain from prompting the respondents' answers too much. Thus, careful and extensive planning, and advice from multiple consultants was used to ensure that the cognitive materials provided would not "lead" the respondents to give certain answers merely because they felt that was what the Interviewer expected.

The CATI dosimetry system developed for use in the HTDS was quite successful in providing a relatively smooth process for the interviews. The staff was successful in identifying appropriate

respondents, and completing interviews when a respondent was available. While there is no way to check the accuracy of the data elicited, the interviewers felt confident that most respondents were able to give responses which accurately reflected their recollections of the experience of the study participants.

E. Scheduling

E.1. Background

Prior to the initiation of the study it was believed that study participants would be widely distributed, with those who had moved away from eastern Washington living primarily in major urban centers in the West and throughout the country. The clinic location sites proposed in the study protocol, therefore, included sites throughout the Pacific Northwest as well as thirteen urban areas across the United States. Once the tracing component of the study began, however, it quickly became apparent that the majority of potential study participants lived in the Pacific Northwest. This made it possible to plan to hold almost all of the clinics within Washington State, with many participants driving to the clinic nearest their home. Those living outside the state could then be flown to Seattle to the clinics held at the FHCRC. Thus, three primary clinic sites were selected to accommodate the majority of study participants: Seattle, Pasco, and Spokane. Additional clinic sites in Walla Walla, Yakima, Wenatchee, and Omak were planned so participants living in these areas would not have to travel as far to attend a clinic. Although one two-day clinic was held in Portland, Oregon early in the study, subsequent clinics for Oregon residents were held in nearby Vancouver, Washington.

There were several advantages to being able to hold nearly all of the clinics within Washington State. First, HTDS could offer all participants a number of different clinic locations in the Northwest. If one location was not convenient, there were others, also relatively close. Second, clinic directions and maps to participants did not have to be constantly re-developed, and there was less potential for error in communicating directions to the participants. Third, many participants located outside the Northwest found the city of Seattle to be an excellent choice for a vacation, and planned their clinic visit to coincide with their vacation plans. Scheduling participants from out of state was also easier. Rather than waiting for all out-of-state participants to be located before scheduling clinics in other regions, they could be brought to Seattle throughout the study, or to other clinic sites, if that was desirable. In addition, holding all clinics in Washington State assisted the blinding of HTDS physicians to residence histories of participants, as it could not be assumed that those living in other states had been less exposed.

The timing and distribution of clinics was determined jointly by the Participation Coordinator and Field Operations Supervisor. As cohort members were located by the Tracing staff, recruiting and CATI efforts were focused so that pools of potential participants for a clinic would be large enough to support full clinic operations. In this way, the clinics could be scheduled at or close to capacity, and more clinics could be scheduled in areas with larger numbers of participants recruited.

A policy was established to provide reimbursement and offer assistance with arrangements for a number of special needs: 1) foreign language interpreters for non-English speaking participants; 2) sign language interpreters for the hearing impaired; 3) personal assistant or companion for participants with a physical or cognitive impairment; and 4) security assistance for participants incarcerated within Washington State. Other special needs were assessed as necessary, and decisions made on a case-by-case basis.

E.2. Objectives of Scheduling

The primary objective of the scheduling activity was to provide each participant with at least three options for clinic attendance, with the least possible inconvenience to the participant. For those participants within driving distance of a clinic, this included providing mileage and meal reimbursement allowances, as well as hotel allowances in the case of overnight trips. For those requiring air travel, all travel arrangements were prepaid by the study and made through FHCRC travel staff, or later, the study's Travel Coordinator, to minimize the inconvenience to participants who had to travel to attend a study clinic.

E.3. Final Process and Procedures

The Clinic Field Operations Supervisor and the Participation Coordinator developed a schedule of clinic dates and locations based on the current residences of participants. The clinic appointment was scheduled after the CATI, or after recruiting, if no CATI respondent was available.

Multiple attempts were made to contact all participants and each participant was offered several options for clinic dates. The Schedulers made calls to participants at varying times of the day and week. All participants, including those scheduled in the final few months of the study, were offered at least three options for clinics. Participants requiring air travel or overnight accommodations were called between 12 and four weeks in advance of the clinic date. Participants not requiring air travel or hotel accommodations could be scheduled up to two weeks before the clinic date.

A computerized tracking system was utilized for tracking the progress of participants through the scheduling process and for creating reports used to generate appointment confirmation letters. Each scheduled participant was sent a letter that included 1) the date and time of clinic appointment; 2) the location of the clinic and directions; 3) travel arrangements summary and/or tickets (if applicable) and 4) the Interview Preparation Worksheet.

If a participant canceled a clinic appointment, the Schedulers attempted to reschedule the participant as soon as possible. A participant who canceled a clinic appointment would be rescheduled an unlimited number of times. If a participant did not show up for a clinic appointment, without notifying the HTDS, the Schedulers attempted to reschedule an appointment. After a participant did not show up for three separate appointment times, no additional attempts were made to schedule the participant. Reminder calls were instituted to reduce the number of "no-shows" at the clinics. These calls were made one to three days prior to the clinic appointment. Based on previous experience with similar epidemiological studies, these reminder calls helped to reduce the number of participants who failed to show for their clinic appointment.

Despite concerted efforts, it was not possible to re-contact some participants after they had agreed to participate (either due to disconnected phone numbers or repeated attempts resulting in no answer or answering machines). In each case, attempts were made to obtain updated information from the CATI respondent (if one was available), through the initial tracing source, or by returning to the tracing staff for further tracing work. If these efforts did not obtain a current telephone number, a letter was sent to the participant requesting they contact us. If attempts to obtain updated information were unsuccessful, or if the participant did not respond to the letters or telephone messages, the participant was classified as "unable to schedule."

The Schedulers assessed the need for travel arrangements and, when necessary, would make the transportation, hotel and other arrangements for the participant. Schedulers followed specific guidelines for allowable travel expenses and reimbursements for participants. The Schedulers completed travel information forms for documenting travel plans.

If a participant decided not to participate in the study during the scheduling process, the Scheduler assessed the reason for the withdrawal and addressed the participant's concerns in an attempt to retain participation. If the participant persisted in the withdrawal, they were asked to complete a Refusal Questionnaire.

If a participant withdrew after agreeing on the first attempt, the decision to re-contact for a second attempt was made by the Scheduler and/or Participation Coordinator, based on the nature of the withdrawal. Second attempts following a withdrawal were handled in the same way regardless of the point at which the withdrawal took place.

E.4. Outcome

E.4.a. Results for the Pilot Study Sample

A total of 1174 Pilot Study Sample participants agreed to participate in the study, and 1063 (90.5%) attended clinics. These figures may differ slightly from those in the Pilot Study Final Report, since efforts to locate, recruit and schedule remaining Pilot Study participants continued throughout the Full Study.

Results in section V.C above, Recruiting, refer to the final agreement status of each participant at the end of the study. It should be noted, however, that some participants actually agreed to participate at the time of recruitment and withdrew from the study at the time of scheduling a clinic appointment. Table V.E-1 shows numbers of those who "Ever Agreed" to participate, those who "Withdrew" from the study prior to being scheduled to attend a clinic, and those who actually attended a clinic, for the Pilot Study Sample.

Table V.E-1. Success in Scheduling Potential Participants - Pilot Study Sample (N=1590)

Scheduling Status	No.	% of Ever Agreed (N=1174)	% of Agreed Excluding Withdrawals (N=1094)	% of Living/Located Pilot Subjects (N=1360)	% of Selected Pilot Subjects (N=1590)
Ever agreed to participate	1174	--	--	86.3	73.8
Withdrew	80	6.8	--	5.9	5.0
Agreed (did not withdraw)	1094	93.2	--	80.4	68.8
Attended clinic	1063	90.5	97.2	78.2	66.9
Unable to schedule*	31	2.6	2.8	2.3	1.9

* Those categorized as "Unable to schedule" are those participants who agreed to participate but attempts to re-contact the participant were unsuccessful. A few additional participants, although offered at least three clinic appointment choices, could not be scheduled before the end of clinics.

E.4.b. Results for the Transition Sample

Table V.E-2 shows numbers of those who "Ever Agreed" to participate, those who "Withdrew" from the study prior to being scheduled to attend a clinic, and those who actually attended a clinic, for the Transition Sample.

Table V.E-2. Success in Scheduling Potential Participants - Transition Sample (N=1005)

Scheduling Status	No.	% of Ever Agreed (N=749)	% of Agreed Excluding Withdrawals (N=692)	% of Living/Located Transition Sample (N=849)	% of Selected Transition Sample (N=1005)
Ever agreed to participate	749	--	--	88.2	74.5
Withdrew	57	7.6	--	6.7	5.7
Agreed (did not withdraw)	692	92.4	--	81.5	68.9
Attended clinic	664	88.7	96.0	78.2	66.1
Unable to schedule*	28	3.7	4.0	3.3	2.8

* Those categorized as "Unable to schedule" are those participants who agreed to participate but attempts to re-contact the participant were unsuccessful. A few additional participants, although offered at least three clinic appointment choices, could not be scheduled before the end of clinics.

E.4.c. Results for the Full Study Sample

Table V.E-3 shows numbers of those who "Ever Agreed" to participate, those who "Withdrew" from the study prior to being scheduled to attend a clinic, and those who actually attended a clinic, for the Full Study Sample.

Table V.E-3. Success in Scheduling Potential Participants - Full Study Sample (N=2604)

Scheduling Status	No.	% of Ever Agreed (N=1939)	% of Agreed Excluding Withdrawals (N=1778)	% of Living/Located Full Subjects (N=2141)	% of Selected Full Subjects (N=2604)
Ever agreed to participate	1939	--	--	90.6	74.5
Withdrew	161	8.3	--	7.5	6.2
Agreed (did not withdraw)	1778	91.7	--	83.0	68.3
Attended clinic	1720	88.7	96.7	80.3	66.1
Unable to schedule*	58	3.0	3.3	2.7	2.2

* Those categorized as "Unable to schedule" are those participants who agreed to participate but attempts to re-contact the participant were unsuccessful. A few additional participants, although offered at least three clinic appointment choices, could not be scheduled before the end of clinics.

E.4.d. Overall Results for the Entire Study

Table V.E-4 shows the final status of scheduling efforts for all those who agreed to participate in the study. Of those who agreed to participate, and did not withdraw from the study at a later time, 96.7% (3447 of 3564) attended a clinic. The rates for withdrawal (7.7%) and for those who did not withdraw but never attended a clinic (3.0%) remained fairly constant throughout the study.

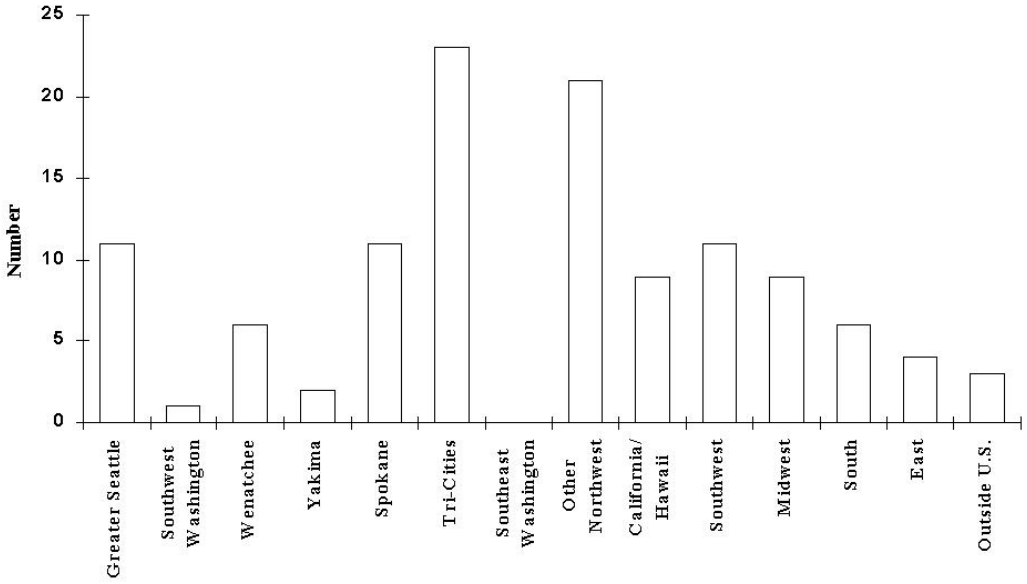
Table V.E-4. Final Success in Scheduling Potential Participants - Entire Study (N=5199)

Scheduling Status	No.	% of Ever Agreed (N=3862)	% of Agreed Excluding Withdrawals (N=3564)	% of Living/Located All Subjects (N=4350)	% of Selected All Subjects (N=5199)
Ever agreed to participate	3862	--	--	88.8	74.3
Withdrew	298	7.7	--	6.9	5.7
Agreed (did not withdraw)	3564	92.3	--	81.9	68.6
Attended clinic	3447	89.3	96.7	79.2	66.3
Unable to schedule*	117	3.0	3.3	2.7	2.3

* Those categorized as "Unable to schedule" are those participants who agreed to participate but attempts to re-contact the participant were unsuccessful. A few additional participants, although offered at least three clinic appointment choices, could not be scheduled before the end of clinics.

While it might be anticipated that those who declined to participate or could not be scheduled would most likely be those participants traveling the longest distances, this did not prove to be the case. It is likely that the popularity of the city of Seattle as a vacation destination and the diligent efforts of the Schedulers to coordinate participant's clinic appointments with their vacation plans, greatly reduced the number who would not attend a clinic due to travel requirements. Figure V.E-1 shows those who agreed to participate but never attended a clinic, by current area of residence.

Figure V.E-1. Subjects Unable to Schedule, by Geographic Area of Current Residence



Note: Southwest, Midwest, South and East regions are defined in section V.C.3.d.

The clinic sites used during the study, number of days at each clinic, and number of participants at each clinic are summarized in Table V.E-5. Approximately 50% of participants attended a clinic in Seattle. Pasco and Spokane were the next most commonly used clinic sites with 21.8% and 13.6% of participants, respectively.

Table V.E-5. Location and Number of Clinic Days and Participants Seen at each Site – Entire Study (N=3447)

Clinic Location	No. of Days at Clinic Site	No. of Participants Seen at Clinic Site	% of Participants
Seattle	133	1719	49.9
Pasco	60	753	21.8
Spokane	37	469	13.6
Vancouver, WA	22	295	8.6
Yakima	7	84	2.4
Walla Walla	4	46	1.3
Portland, OR	2	33	1.0
Wenatchee	1	17	0.5
Omak	1	16	0.5
Colville	2	15	0.4
Total	269	3447	100

The current residences of participants and the clinics they attended are shown in Table V.E-6. While most participants attended the clinic nearest their home, a number of participants attended a clinic in other areas. This occurred either as an incentive to participation (e.g., the participant wanted to travel to visit family) or due to difficulty scheduling the participant at a clinic in their area. The latter reason was most common in the smaller, more rural communities, where fewer clinics were held.

Table V.E-6. Current Residence of Participants by Clinic Site – Entire Study (N=3447)

	Clinic Site Attended								
Current Residence	Seattle	Pasco	Spokane	Vancouver, WA Portland, OR	Yakima	Walla Walla	Colville	Wenatchee	Omak
Greater Seattle	583	3	1	6	1	1	0	0	0
SW Washington	3	3	1	78	0	0	0	0	0
Wenatchee	11	19	23	0	6	0	1	16	15
Yakima	8	50	0	0	76	1	0	1	0
Spokane*	5	6	329	1	0	1	14	0	1
Tri-Cities	6	559	7	1	1	38	0	0	0
SE Washington	0	1	6	0	0	0	0	0	0
Other Northwest	117	80	79	220	0	5	0	0	0
California/Hawaii	318	10	3	10	0	0	0	0	0
Southwest US	237	12	7	3	0	0	0	0	0
Midwest US	164	4	6	4	0	0	0	0	0
Southern US	146	3	2	3	0	0	0	0	0
Eastern US	99	1	3	2	0	0	0	0	0
Out of US	22	2	2	0	0	0	0	0	0

* Includes the Colville area.
Note: Southwest, Midwest, South and East regions are defined in section V.C.3.d.

E.4.e. Conclusions

The process developed for scheduling participants proved to be extremely effective with nearly 90% of those who ever agreed to participate completing a clinic. The number of participants who withdrew after initially agreeing (298) was not unexpectedly high, considering the requirements of the clinical thyroid exam and interview, as well as the amount of travel required by many participants. The Schedulers arranged travel, including airline and/or hotel arrangements, for 1288 (37.4%) of the 3447 participants attending a clinic.

The participants who did not officially withdraw from the study, but who could not be scheduled into a clinic, were offered a minimum of three clinic options, and most were offered many more opportunities. In many cases, those in the category of "unable to schedule" and/or withdrew had multiple reasons for being unable to attend. These reasons included illness, unpredictable work schedules, and family responsibilities. Although some participants withdrew from the study during the scheduling process or were unable to be scheduled because of scheduling conflicts, many initially reluctant participants were persuaded by the scheduling staff to attend a clinic.

F. Clinical Evaluation

F.1. Background

F.1.a. Objectives for Clinical Evaluations

The objective of the clinical evaluation was to provide a thorough clinical examination of each study participant to identify the presence of thyroid disease or primary hyperparathyroidism. The evaluation provided data to determine the current presence or absence of disease for each disease endpoint. In addition, the clinic visit provided an opportunity for participants to ask questions and receive information about radiation and thyroid disease.

F.1.b. Rationale

The clinical evaluation provided information on each participant's current thyroid and parathyroid disease status. Prior to the examinations, each participant was asked to respond to an In-Person Interview (see section V.G.) which included questions pertaining to history of thyroid disease or hyperparathyroidism. The clinical evaluation included a thyroid ultrasound scan, independent thyroid examinations (palpation) by two physicians specializing in thyroid disease, and blood collection for thyroid function, parathyroid function, and anti-thyroid immune response tests. Additional studies were requested if indicated by the presence of palpable thyroid nodules. All costs associated with the clinical work-up, as well as travel to and from the clinic sites, were paid by the study.

F.2. Clinic Procedures

F.2.a Clinic Locations and Schedules

All clinics were held in Washington State, except for one 2-day clinic in Portland, Oregon. Conducting clinics in Washington provided efficiencies in equipment transport, planning, set-up, and staff travel. The clinics at the Fred Hutchinson Cancer Research Center in Seattle were primarily held on Friday and Saturday. Suitable clinic space was usually not available on Friday in other locations so clinics were most often held on Saturday and Sunday in locations other than Seattle.

F.2.b. Clinical Evaluation Process

Specific procedures for clinic operations were developed to optimize efficiency, assure all steps were completed, minimize waiting and maintain confidentiality. A packet was prepared for each scheduled participant containing consent forms and the data forms to be completed at the clinic. Participant names were written on index cards that were removed from the packets and transported separately for purposes of confidentiality and then re-attached to the packet at the clinic site. A clinic flow sheet was attached to the front of the packet. The clinic flow sheet provided an outline of each step or station to be visited and included a list of all data forms. The clinic staff person performing each clinic activity would check-off the completed activity.

Participants would first check in with the clinic coordinator and sign the study consent form (Appendix 12). The participant was then escorted to the In-Person Interview. After the interview, he or she was taken to the blood draw station. Following the blood collection, the nurse or phlebotomist answered any participant questions.

The next step was the ultrasound scan. The ultrasound scan was recorded on videotape, prints were made of key findings in the exam, and the sonographer completed a Thyroid Ultrasound Form

(Appendix 13). Following the ultrasound scan, thyroid exams were conducted separately by each of the two physicians at the clinic and the results recorded on Thyroid Examination Forms (Appendix 14). After the two independent exams, the physicians would confer. If there was disagreement between the examiners, they would perform another thyroid exam together to reach a consensus and complete a Consensus Examination Form (Appendix 15). After the two physicians reached an agreement on the exam findings and recorded the results, they reviewed the ultrasound scan results and the ultrasound scan prints. If there was any disagreement between the ultrasound scan and the physical exam results, a post-ultrasound consensus exam was done together by the two physicians and a Post-Ultrasound Consensus Examination Form completed. The physicians then conferred privately and discussed the results of the exams and the ultrasound scan. The physicians returned to the exam room to discuss the exam and ultrasound findings with the participant. A Thyroid Ultrasound Fact Sheet was given to all participants (Appendix 16).

If a fine needle aspiration (FNA) was indicated as a result of the exam or ultrasound scan, the physicians would discuss this recommendation with the participant and request consent to perform the procedure. FNA procedures were performed at any time throughout the day according to the participant's schedule or request.

The final step was a check-out with the Clinic Coordinator. The coordinator reviewed the clinic flow sheet to be certain that all steps were completed and checked that all data forms were completed. Travel reimbursement paperwork was completed at check-out.

At the end of the last clinic day, staff packed up all clinic supplies and equipment. Participant names were removed from the individual packets and the packets were transported to the HTDS office in a locked suitcase. Clinic staff transported serum specimens to the Pacific Medical Center Laboratory in Seattle and transported FNA specimens to the Laboratory of Pathology at Swedish Hospital in Seattle. Ultrasound videotapes were sent by messenger to Seattle Nuclear Medicine Associates.

F.2.c. Clinic Staffing

Clinic staff consisted of two physicians, one nurse or phlebotomist, one ultrasonographer, two or three interviewers, and a Field Operations Coordinator or Supervisor. Approximately fifteen potential participants were evaluated on each day. The dates and locations of clinics and staffing assignments were finalized 3-4 months in advance. To improve efficiency at clinics, various members of the clinic staff were trained and capable of performing multiple tasks. The Field Operation Supervisor was a certified phlebotomist and was trained to assist with FNA procedures. Both Field Operation Coordinators were trained in the In-Person Interview and one was also able to assist with FNA procedures. In addition, one interviewer was a certified phlebotomist, allowing her to assist the nurse or phelbotomist at peak times, or move to this position if needed.

F.2.d. Efforts to Reduce Physician and Ultrasonographer Bias

To ensure that the clinical decisions by the physicians and sonographers were not influenced by knowledge of the participants' possible exposures to Hanford ^{131}I, several precautions were taken during the clinical evaluation. The nurse or the phlebotomist asked participants not to speak with the physicians and sonographers about where they had lived, or about the possibility of their exposure to radiation from Hanford. Signs were also posted throughout the clinic requesting that participants not discuss these issues with physicians and sonographers. In addition, some participants who lived in towns where clinics were held were asked to attend clinics in other areas, so that physicians and sonographers would not associate participants at one clinic site with exposure and those at another clinic site with non-exposure. Finally, physicans were required to record at the end of their evaluation of each participant whether he had any indication fo possible radiation exposure for that individual. Of a total of 3440 evaluable participants,

there were only 15 instances where the physician had some suspicion that the participant might have had prior radiation exposure.

F.3. Serum Sample

F.3.a. Laboratory Studies

The Research Nurse collected a blood sample for thyroid function and other laboratory studies. A small number of persons refused to provide a blood sample. Such refusal did not affect the participant's eligibility for participation in the study or evaluability (see section IV.B above). Three 10 cc tubes of blood were drawn and centrifuged on site. The serum was transported within 72 hours to Seattle where one tube was frozen at -70 degrees Centigrade and stored as a reserve. The remaining two tubes were transported to the clinical laboratory at Pacific Medical Center in Seattle for the following studies:

- TSH (Thyroid Stimulating Hormone)
- FTI (Free Thyroxine Index)
- Antithyroid Antibodies
- Calcium

F.3.b. Changes in Laboratory Assays

F.3.b.1. AMA to Anti-TPO

Specific tests and assays changed throughout the course of the study, prompted by changes in the industry standard and on the recommendations from the laboratories.

The antimicrosomal antibody (AMA) assay was used initially to screen for autoimmune thyroid disease. Due to improvements in laboratory assays, the anti-thryroperoxidase (Anti-TPO) assay was available from Pacific Medical Center Laboratory in September 1995. At the request of the HTDS, the two assays were run in tandem until more than 500 assays had been performed using both methods. An analysis was performed to ensure the two methods were comparable, after which, the AMA was discontinued. The results of this analysis are shown in Table V.F-1 below.

Table V.F-1. Agreement between AMA and Anti-TPO Assay Results (N=677)

		Anti-TPO		
		Negative	Positive	Total
AMA	Negative	480	19	499 (73.7%)
	Positive	49	129	178 (26.3%)
	Total	529	148	677

These results indicate a high level of agreement (90%) between the two assay methods.

F.3.b.2. TSH Methods - RIA, EIA-1, EIA-2

The TSH test methods performed on HTDS serum specimens were done initially by radioimmune assay (RIA). The RIA method was used from November 1992 through January 1994. The RIA TSH method was changed to an ELISA immunometric assay (designated EIA-1) method starting February 1994. The EIA-1 method was used from February 1994 through August 1995. The EIA-1 method was modified to the EIA-2 method in September 1995. The EIA-2 method was used from September 1995 until the end

of the study in September 1997. In addition, the TSH normal range from the EIA-2 method was changed from a range of 0.47-5.01 units to 0.32-5.01 units as of Jan. 10, 1997.

F.3.b.3. Parathyroid Hormone Methods

Measurement of Intact PTH was done for all participants with an elevated serum calcium level. From the first clinic in November 1992 through October 1994, the Intact PTH test was done by the immunoradiometric assay. From November 1994 through the last clinic in September 1997, PTH was done by two methods, the IRMA and the Chemiluminescence methods. Separate calcium levels accompanied each method.

F.3.b.4. Anti-TG

In 1998 after the clinics were completed, anti-thyroglobulin antibody (anti-TG) assays were performed on serum samples that had been frozen and stored from the blood samples provided by HTDS participants at the study clinics. Although the anti-TPO antibody served as the highest quality assay for autoimmune thyroid disease, recent improvements in the anti-TG assay were available through Dr. Carole Spencer, an international expert in the measurement of antithyroid antibodies. These assays, which were performed in Dr. Spencer's laboratory, provided an opportunity to assess more fully the cumulative incidence of autoimmune thyroiditis in the HTDS cohort.

F.4. Inclusion of an Ultrasound Exam

The clinical evaluation included a thyroid ultrasound scan to detect thyroid nodularity. The decision to include an ultrasound scan in the clinical evaluation was based on three primary benefits: 1) there would likely be a small increase in the ability of the study to detect a radiation effect associated with clinical thyroid disease as currently defined; 2) nonpalpable thyroid UDAs abnormalities of the thyroid could be included as a study outcome variable and 3) the recorded ultrasound scan provided an objective record of the presence, location, and characteristics of thyroid growth abnormalities.

A certified ultrasound technologist performed the scan and was blinded to the participants' exposure status. The entire thyroid ultrasound scan was recorded on videotape. Physicians examined the participant without any knowledge of the ultrasound findings, then again after viewing printouts from the scan. Following the clinic, the videotaped scans were transported to Seattle for review by an off-site radiologist.

F.5. Ultrasound Follow-up Program

F.5.a. Purpose of the Ultrasound Follow-up Program

Participants at clinics who were found to have nonpalpable thyroid abnormalities seen only on the ultrasound scans were given a Thyroid Ultrasound Fact Sheet. This fact sheet explained the unknown clinical significance of the abnormal findings. In addition, these participants were invited to participate in the HTDS Ultrasound Follow-Up Program.

The Ultrasound Follow-Up Program was offered as a service to participants, and was not intended as a substitute for treatment or follow-up by participants' health care providers. The primary purpose of the program was to: 1) identify early nonpalpable, rapidly growing, thyroid cancers, and 2) provide referral assistance to facilitate appropriate management of participants' medical care.

Initially, there were two possible follow-up appointments for eligible participants. The first appointment was at 9 months after the initial clinic appointment and, if a change was detected on physical exam or ultrasound scan, a second follow-up appointment was recommended 6 months after that date. This second exam was a total of 15 months after the participant's initial clinic appointment. This design was modified in January 1994 to become a one-time follow-up appointment done 9-15 months after the participant's initial clinic appointment.

The original design of the follow-up program also included an examination by an HTDS physician. In February 1994, the physician examination was discontinued as part of the follow-up program. The purpose of the follow-up program was to detect changes in nonpalpable thyroid cancers. Since it was very unlikely that small changes in size could be detected by physical examination, it was decided that little useful information was provided by the follow-up physical exam. If a new or larger nodule was found on follow-up ultrasound exam, the participant was examined by a physician. A total of 260 participants were evaluated during the Ultrasound Follow-Up Program.

F.5.b. Discontinuation of the Ultrasound Follow-up Program

The Ultrasound Follow-Up Program was discontinued in June of 1995 for several reasons. In May 1995, a physician review of the data collected from the follow-up program revealed that no significant changes were found between the initial and follow-up ultrasound scans that would change the diagnosis or the recommended treatment or follow-up. Consequently, very little new diagnostic information had been collected from the follow-up program and no fast-growing cancers had been identified. The Ultrasound Follow-Up Program was not one of the HTDS research objectives and data from the follow-up exams and scan were not entered into the primary database. Therefore, discontinuation of the follow-up program did not affect the study's objectives.

The follow-up program became difficult to integrate into the busy HTDS clinic schedule. The follow-up program utilized the same ultrasound equipment and personnel as the HTDS clinics. An assessment of the clinic schedules indicated that the follow-up program would cause a significant delay in the completion of the HTDS clinical evaluations. Also, scheduling of the follow-up clinics was determining the dates and locations of HTDS clinics rather than consideration of new participants' residences.

An additional operational concern was the volume of work generated by the follow-up program. The Ultrasound Follow-up Program demanded substantial staff time and effort at clinics, and in the study office for follow-up appointment calls and letters, entry of tracking data, and physician review of the results and preparation of follow-up outcome letters to participants and their personal physicians. Continuation of the follow-up program would have required hiring additional staff and purchasing additional equipment.

In June of 1995, after consultation with the CDC and the HTDS Advisory Committee, the ultrasound follow-up program was discontinued. A special fact sheet was developed for health care providers and participants that provided information on the significance and management of patients with nonpalpable thyroid ultrasound detected abnormalities (UDAs).

F.6. Physicians

The study began with two HTDS physicians. In April 1993, four physicians were added to meet the demands of the full clinic schedule. All physicians were thyroid specialists. Physician pairings and clinic locations were rotated among physicians to reduce the potential for bias that might occur if the same physicians worked only at certain clinic locations.

A total of 3447 eligible participants attended a clinic, however, one participant did not have a thyroid exam due to a tracheotomy. Three of the 3446 participants who had a thyroid exam were examined by one physician, rather than two because of a scheduling problem. The numbers of participants examined by each pair of physicians is shown in Table V.F-2. Physicians #1 and #2 participated from the beginning of the clinical activity and continued throughout the study. As a result, 746 (21.6%) of the participants were examined by physicians #1 and #2, and 2822 (81.9%) were examined by a physician pair that included physician #1 and/or #2. The three participants examined by a single physician were all seen by physician #1 or #2.

Table V.F-2. Pairings of Physicians for Clinical Examinations*

		Second Physician				
First Physician	No Second Physician	#2	#3	#4	#5	#6
#1	1	746	285	322	367	198
	0.03%	21.6%	8.3%	9.3%	10.7%	5.7%
#2	2	--	332	423	66	82
	0.06%	--	9.6%	12.3%	1.9%	2.4%
#3	--	--	--	117	67	233
	--	--	--	3.4%	1.9%	6.8%
#4	--	--	--	--	26	141
	--	--	--	--	0.8%	4.1%
#5	--	--	--	--	--	38
	--	--	--	--	--	1.1%

* Entries in the table are the number (upper) and percentage of participants who attended the clinic and were examined by the indicated pair of physicians.

F.7. FNA Criteria

The original study protocol called for FNA procedures to be performed on study participants whose exams indicated the presence of discrete, palpable, solitary thyroid nodules or discrete, dominant nodules in a multinodular thyroid gland.

In February 1994 the criteria for conducting FNA procedures at clinics were expanded. In addition to nodules palpated on exam, the HTDS physicians also requested consent to perform FNA on participants who were found to have nonpalpable ultrasound detected nodules of 1.5 cm or greater (average of three dimensions) in a palpable thyroid gland. This modification was made after several participants were found to have quite large abnormalities detected by ultrasound that neither of the two experienced thyroidologists could palpate. The decision to attempt to perform an FNA on these large, ultrasound detected thyroid abnormalities was based on: 1) consideration of the HTDS physician's confidence of biopsying the nodule(s) detected by the ultrasound; 2) the physician's concern that the abnormality may represent a thyroid neoplasm; and 3) technical and safety aspects of performing a biopsy on a nonpalpable abnormality.

In some cases, the HTDS physicians recommended an FNA to a study participant after his/her clinic appointment. This recommendation was made as a result of the radiologist's review of a participant's ultrasound scan results.

In a very few cases, participants were recommended to undergo ultrasound-guided FNA as a safety precaution due to a nodule's close proximity to the carotid artery. In these rare instances, the participant was referred to a medical facility near their place of residence that had the capability for ultrasound-guided FNA. The data collected from these procedures were used in the data analysis.

After review of the FNA biopsy results, each participant was assigned a diagnosis from the FNA or was recommended to have further evaluation. If the FNA was consistent with thyroid cancer (papillary carcinoma), the participant was recommended to see his/her physician for consideration of thyroid surgery. As discussed in section V.I below, all histology slides from such surgery were requested for review by the HTDS study pathologist and the diagnosis of either thyroid cancer or benign thyroid nodule was assigned depending on the pathology review. If the HTDS FNA result was adequate and consistent with a benign thyroid nodule, no further evaluation of the nodule was recommended, and the participant was recommended to follow-up with their personal physician. In these cases, the HTDS diagnosis was benign thyroid nodule.

If the initial FNA biopsy result indicated an intermediate or high probability of a follicular neoplasm (either benign or malignant), the participant was recommended to have further evaluation by his/her physician team, usually with consideration of thyroid surgery. These recommendations were made since FNA cytology cannot reliably distinguish between a benign follicular neoplasm (adenoma) and a follicular carcinoma. In these participants, no HTDS diagnosis was initially assigned but rather the participant was followed until the end of the study to await further diagnostic information, usually from surgery. When such information became available, the participant was then given an HTDS diagnosis of either thyroid cancer or a benign thyroid nodule depending on the outcome of the surgical diagnosis. For participants who (for whatever reason) did not go on to have thyroid surgery by the end of the HTDS field component (1997), definitive information to make a diagnosis on the nodule that was biopsied was not available. For the HTDS analysis, these individuals were classified as having a nodule "suspicious for follicular neoplasm". It is important to emphasize that none of the participants with this diagnosis had a nodule that was suspicious for papillary thyroid cancer but rather a nodule that had some probability of a follicular neoplasm. Since the majority of such lesions represent benign follicular adenomas, this category largely would be expected to represent benign nodular lesions. The following data from the HTDS illustrate this further.

Of the 259 evaluable participants who underwent FNA, 47 (18.1%) were recommended to have further biopsy or surgery. Of these 47, 12 were subsequently found to have thyroid cancer, five to have follicular adenoma, and 13 to have benign nodule other than follicular adenoma. The remaining 17 participants (6.6% of the original 259) were classified as suspicious for follicular neoplasm. All of these 17 cases were so classified because they did not go on to have further biopsy or surgery. For none of these 17 participants was there an actual clinical suspicion of papillary cancer. In fact all 17 had intermediate or high probability of follicular neoplasm based on their FNA results. Thus, for the 6.6% of the 259 persons who had FNA and were recommended to have further biopsy or surgery, we were not able to obtain further cytological or histological diagnoses. While the absence of such diagnoses makes it impossible to rule out the possibility of thyroid cancer, the probability of a benign lesion would be quite high given that all the 17 cases were suspicious for follicular neoplasm rather than for papillary carcinoma.

F.8. Thyroid Nuclear Scan Criteria

A thyroid nuclear scan and radioiodine uptake was recommended for three situations: 1) the results of an FNA indicated suspicious cytology which could be an indicator of an autonomously functioning nodule; 2) a neck mass was felt in the physical exam which was suggestive of an abnormality, but because of a technically difficult exam (e.g., a very obese neck), a consensus between examiners could not be reached at the clinic; and 3) for participants who had a suppressed TSH blood value and a normal or elevated FTI blood value to rule out a diagnosis of Graves Disease or a toxic thyroid nodule.

F.9. Training and Quality Control

F.9.a. Training

Two months prior to the first HTDS clinic, a "mock" clinic was held for staff training. HTDS staff assumed the roles of study participants and went through each clinic activity including the interview, blood draw, thyroid ultrasound scan and physical exam. Blood specimens were sent to the laboratory and analyzed to test specimen processing, transport, and other clinic procedures.

The mock clinic accomplished three primary objectives. First, the anticipated amount of time a participant would spend at each activity, and the total time at the HTDS clinic were verified to be consistent with the predicted total of two hours. The second objective was to test the designed clinic flow. The goal was to ensure the smooth and orderly transfer of participants through the various steps of the clinics, to avoid long waits, and to assure that each activity would be completed. The third objective was to give HTDS staff firsthand experience of the clinic activities to raise their awareness and ability to respond to participant questions and concerns about any part of the clinic experience.

The study ultrasonographers underwent additional training with Dr. Keith Wang of Seattle Nuclear Medicine Associates to standardize their technique of performing thyroid ultrasound scans. New sonographers were accompanied by the experienced HTDS sonographers for a minimum of two full days or until agreement in technique was obtained. During this training period, one sonographer performed the exam and recorded the results while the second sonographer recorded the findings on a second Thyroid Ultrasound form. The sonographers then switched places for the next participant. At the end of the clinic, the Field Operation Supervisor compared the findings for each participant, reviewed any discrepancies with both sonographers, and instituted further training as necessary.

F.9.b. Ultrasonographer Quality Control

A total of four certified sonographers worked on the study at various times but only one or two sonographers were on staff at any given time. An attempt was made to divide the clinical schedule evenly between the two sonographers on staff. In addition to the initial training, ongoing quality control procedures were undertaken to monitor inter-operator reliability. Approximately every two months, both sonographers would perform independent scans on each of five participants. The results of the scans were recorded on separate videotapes, the findings were compared and discrepancies were noted and discussed. Quality of sonography outcomes was monitored for each pair of sonographers that were currently sharing clinics. Results of the ultrasound quality control comparisons, based on a total of 103 participants, are summarized in Table V.F-3.

Table V.F-3. Results of Quality Control Ultrasound Studies

		A	B		No. of Participants		
Ultrasound Tech Pair	No. of Participants	No. of Nodules (>5 mm) Identified by Either Tech	No. of Nodules (>5 mm) Identified by Both Techs	B, as a Percent of A	With No Nodules (>5 mm) by Both Techs	With No Nodules (>5 mm) by Tech 1 but ≥1 Such Nodules by Other Tech	With >1 Nodule (5 mm) by Tech 1, but No Such Nodules by Other Tech
1 + 2	30	33	20	61%	16	1	1
1 + 3	25	27	12	44%	16	1	2
1 + 4	48	79	44	55%	23	1	4

F.9.c. Radiology Quality Control Program

A radiology quality control program was designed to monitor agreement rates among the radiologists interpreting the videotaped ultrasound scans. Six radiologists were initially identified to review and interpret the ultrasound videotapes for the HTDS. The radiologist assigned to read ultrasound scans for a particular clinic was determined solely by the radiologists' work schedules and availability. No effort was made to equalize the numbers of scans read by each radiologist. One radiologist interpreted the scans from the first clinic only, and one radiologist read scans only through the second month of the study. From early 1993 until the end of the study, four radiologists were involved in interpreting the HTDS ultrasound scans.

For purposes of quality control, approximately ten scans per month were sent back to the radiologists to be reviewed and interpreted a second time. These tapes were submitted along with scans from the most recent clinic. Comparisons between the two forms of abstracted findings by the radiologists were made to determine if significant changes could be identified between the first and second reading. A total of 343 ultrasound exams were interpreted twice. In most cases the second review was performed by a radiologist other than the one who originally reviewed the case. However, in a few cases, the quality control review was done by the radiologist who first reviewed the case due to the radiologists' scheduling. As shown in Tables V.F-4 through V.F-7, there were very high levels of concordance between the results of the original reviews and second review.

Table V.F-4. Radiologist Agreement on Presence of Any Nodule

	QC Radiologist			
Clinic Radiologist	Yes	No	Uncertain	Total
Yes	132	6	2	140
No	4	195	2	201
Uncertain	1	0	1	2
Total	137	201	5	343

Table V.F-5. Radiologist Agreement on Number of Nodules Less Than 5mm Average Dimension

	QC Radiologist			
Clinic Radiologist	0	<10	˜10	Total
0	289	4	0	293
<10	7	40	0	47
˜10	2	0	1	3
Total	298	44	1	343

Table V.F-6. Radiologist Agreement on Presence of Diffuse Abnormalities

	QC Radiologist		
Clinic Radiologist	Yes	No	Total
Yes	30	8	38
No	7	298	305
Total	37	306	343

Table V.F-7. Radiologist Agreement on Number of Nodules ≥5mm Average Dimensions

		QC Radiologist					
		0	1	2	3	˜3	Total
Clinic Radiologist	0	220	-	-	-	-	220
	1	-	69	-	-	-	69
	2	-	1	25	-	-	26
	3	-	-	-	9	1	10
	˜3	-	-	-	-	18	18
	Total	220	70	25	9	19	343

F.10. Outcome and Results

F.10.a. Results for the Pilot Study Sample

Table V.F-8. shows the number of Pilot Study participants completing each component of the clinic. A total of 1063 Pilot Study participants attended a clinic. All except four participants (99.6%) had blood drawn for thyroid function and other studies. Seventy-six of the 79 participants for whom fine-needle aspiration was recommended had the procedure performed. This represents a 96.2% consent rate for FNA, significantly higher than had been anticipated.

Table V.F-8. Summary of Clinic Participation - Pilot Study Sample

	Clinic Components Completed	
	No.	% [A]
Agreed to participate	1094	--
Attended clinic	1063	100
In-Person Interview	1063	100
Ultrasound examination	1063	100
Radiologist review of ultrasound	1063	100
Blood sample drawn	1059	99.6
All thyroid function tests performed and results obtained	1058	99.5
Thyroid examination by two physicians	1061	99.8
Thyroid examination by one physician	2	0.2

[A] Percentage calculated in relation to number who attended clinic.

F.10.b. Results for the Transition Sample

Table V.F-9 shows the number of Transition participants completing each component of the clinic. A total of 664 participants from the Transition Sample attended a clinic. All except one (99.8%) had blood drawn for thyroid function studies. Forty-three (97.7%) of the 44 for whom FNA was recommended had the procedure performed.

Table V.F-9. Summary of Clinic Participation - Transition Sample

	Clinic Components Completed	
	No.	% [A]
Agreed to participate	692	--
Attended clinic	664	100
In-Person Interview	664	100
Ultrasound examination	664	100
Radiologist review of ultrasound	664	100
Blood sample drawn	663	99.8
All thyroid function tests performed and results obtained	661	99.5
Thyroid examination by two physicians	663	99.8
Thyroid examination by one physician	1	0.2

[A] Percentage calculated in relation to number who attended clinic.

F.10.c. Results for the Full Study Sample

Table V.F-10 shows the number of Full Study participants completing each component of the clinic. A total of 1720 participants from the Full Study sample attended a clinic. All except four (99.8%) had blood drawn for thyroid function studies. Of the 149 for whom FNA was recommended, 140 (94.0%) had the procedure performed.

Table V.F-10. Summary of Clinic Participation - Full Study Sample

	Clinic Components Completed	
	No.	% [A]
Agreed to participate	1778	--
Attended clinic	1720	100
In-Person Interview	1720	100
Ultrasound examination	1719	99.9
Radiologist review of ultrasound	1719	99.9
Blood sample drawn	1717	99.8
All thyroid function tests performed and results obtained	1713	99.6
Thyroid examination by two physicians	1719	99.9
Thyroid examination by one physician	0	0

[A] Percentage calculated in relation to number who attended clinic.

F.10.d. Overall Results for the Entire Study

Table V.F-11 shows the number of participants from the entire study completing each component of the clinic. A total of 3447 eligible participants attended an HTDS clinic. Seven of these participants were judged non-evaluable (see section IV-B for definition of evaluable participant) following their clinic participation, one due to inability to perform a thyroid exam due to a tracheotomy, and six because of incomplete residence histories. Of the 3447 participants, 3439 (99.8%) had blood drawn for thyroid function studies, and 3446 had a thyroid ultrasound scan. Three participants were examined by only one physician due to scheduling difficulties. Of the 272 participants for whom FNA was recommended, 259 (95.2%) underwent the procedure, while 28 (96.6%) of the 29 participants recommended to have a nuclear scan complied.

Table V.F-11. Final Summary of Clinic Participation - Entire Study

	Clinic Components Completed	
	No.	% [A]
Agreed to participate	3564	--
Attended clinic	3447	100
In-Person Interview	3447	100
Ultrasound examination	3446	99.97
Radiologist review of ultrasound	3446	99.97
Blood sample drawn	3439	99.8
All thyroid function tests performed and results obtained	3432	99.6
Thyroid examination by two physicians	3443	99.9
Thyroid examination by one physician	3	0.1

[A] Percentage calculated in relation to number who attended clinic.

F.10.e. Conclusions

One indication of the success of the HTDS clinics is the excellent overall completion rates for each component of the clinical evaluation, particularly the FNA procedures. An emphasis was placed on establishing a caring and supportive environment for participants and reducing the level of stress to participants during the medical examinations.

G. In-Person Interview

G.1. Background

The standard In-Person Interview (IPI) consisted of questions designed to collect information about the following areas: 1) residences after age 15 to identify participants who may have received radiation exposure living near other nuclear facilities; 2) occupational history, to account for possible on-the-job radiation exposure; 3) smoking history; 4) medical and dental radiological procedures or radiation therapy after age 15 to complete the identification of radiation exposure to the thyroid from these sources begun in the CATI; 5) thyroid disorders after age 15 to complete the medical history begun in the CATI; 6) prescription drug history, to identify those persons whose thyroid disease may be a side effect of certain prescription medications, or who are now taking medications which could impact the results of thyroid assays performed at the clinic; 7) standard demographic questions; and 8) familiarity/bias questions to determine if a relationship exists between the answers given in the questionnaire, and the participant's knowledge or beliefs about the Hanford radiation releases.

The questions on the standard IPI (Appendix 17) covered the time period beginning after age 15 and extending to the present because detailed information about the subject from birth through age 15 was obtained in the CATI. However, since it was anticipated that a CATI respondent might not be available for all potential participants, an expanded version of the IPI (Appendix 18) was designed to also collect information from birth through age 15 that would have been provided by the CATI (see section V-D for more information regarding CATI).

G.2. Objectives of the In-Person Interview

The primary purpose of the IPI was to obtain information directly from the study participant about past occupational or medical radiation exposures, history of thyroid disease, and general demographic information. Most questions in the standard IPI pertained to the period after age 15 to the present because the CATI provided information about the period from birth through age 15. Participants who did not have a CATI were given an expanded version of the IPI for collecting key data for the period from birth to age 15. This expanded version of the IPI provided details about residence history and types of milk consumed which were necessary to estimate a Hanford radiation dose. The IPI was conducted before the participant completed the medical components of the thyroid clinical evaluation (ultrasound, blood draw, and physical examination) to ensure that the participant's responses would not be influenced by knowledge of exam results.

G.3. Development and Revision of the Questionnaire

A total of six versions of the standard and expanded In-Person Interviews were used in the three phases of the study. With the exception of a modification of the residence history questions following the Pilot Study, the differences between versions consisted of minor wording changes made for clarification purposes and deletion of questions determined to be unnecessary. Listed below is a summary of the revisions:

November 4, 1992	Original version
January 6, 1993	Wording changes for clarification; income categories in demographics section adjusted; change in the order of questions in the prescription drug section
December 20, 1994	Pilot Study Revisions: Information about residences asked after 1957 only for geographic areas near other nuclear production facilities or test sites; mother's residence history while pregnant

	was added to the expanded version for participants born after December 1944; questions on whether the participant had ever been diagnosed with hyperparathyroidism were added
June 28, 1995	Deleted questions about other names used by the participant, except those for whom historical medical records were being sought; deleted question on reasons why the participant thought they didn't know more about Hanford; minor wording changes in two areas of participant directions
December 11, 1995	Revised wording from "x-ray treatment" to "radiation treatment" in medical history section; modified the explanation of fluoroscopy for clarification

G.4. Procedures for the In-Person Interview

All interviews were conducted by trained and experienced Interviewers at the time of the participant's visit to the HTDS clinic. The interview was always completed before the thyroid examination to eliminate the possibility that the participant's answers to the interview questions may be influenced by the results of the thyroid exam. Prior to the initiation of the interview, each participant was required to read and sign a consent form agreeing to participate in the study.

An In-Person Interview Preparation Worksheet (Appendix 19) was sent to the participant two weeks before the clinic appointment. Participants were asked to complete the worksheet prior to attending the HTDS clinic, and to refer to this form during the interview. At the end of the interview, the worksheet was collected from the participant to be filed with the questionnaire.

Following each In-Person Interview, the interviewer recorded his or her subjective impression of the reliability of the data collected (High, Generally Reliable, Questionable, Unreliable) and the participant's level of cooperation (Very Good, Good, Fair, Poor).

Interviewers were not assigned to any particular clinics or counties, thereby reducing the potential for bias that might occur if area-specific assignments were made.

G.5. Training and Quality Control

Interviewers were initially trained by the same Field Operations Coordinator and experienced FHCRC interviewing personnel to assure uniformity and quality in the interviewing procedures. Training consisted of instruction in general interviewing skills, proper methods and timing of probing, detailed question-by-question instruction, and instruction on editing and callbacks. Training sessions included role-playing exercises. Interviewers pilot tested the questionnaires and worksheets and refined their skills by interviewing a small sample of volunteers. Training was supplemented by two manuals: 1) a general Interviewing Manual and 2) a Question-by-Question Manual for the standard and expanded versions of the In-Person Interview.

The Interviewers edited (reviewed) each questionnaire at the clinic site immediately after the interview was completed to assure all information was completely filled out and to identify discrepancies. The Field Operations Coordinator edited the interview a second time (over-edited) within 14 days of the date the interview was conducted. Re-contacting of study participants by the Field Operations Coordinator for clarification or missed questions was usually done within two weeks of the date of the original interview. The Field Operations Coordinator coded the questionnaires for data entry at the time of over-editing. A manual for coding of interviews was developed and documents the coding procedures utilized.

Call-backs to participants for clarification or additional information were limited to those instances where the HTDS Interviewer made an error, either by omitting a question or not adequately probing a question. Most decisions on whether to call back a participant for additional information relating to the residence history were discussed with a study investigator before contacting a participant to determine whether information obtained after the thyroid examination could be used.

G.6. Outcome and Results

At the conclusion of the study, a total of 3447 eligible participants had attended the HTDS clinic. No participants declined to complete an In-Person Interview. Review of the interviews resulted in identifying six questionnaires judged to have insufficient residence history information to calculate a dose estimate. These six participants were determined to be non-evaluable (see section IV.B for definition of evaluable participant). One participant was unable to complete the interview because of developmental disabilities, however the participant's father (who was unable due to illness to participate in a CATI) was mailed a modified version of the expanded interview and provided the dosimetry and In-Person Interview information in this manner. Some participants with developmental or other disabilities were accompanied during the interview by a family member or guardian, who aided in the interview process.

Table V.G-1 is a summary of standard and expanded interviews completed during each phase of the study. The passage of time and selection of participants from earlier birth years later in the study increased the use of the Expanded In-Person Interview in the Transition and Full Study Samples. Overall, 61% of participants completed the Standard In-Person Interview, while 39% completed the expanded version. The 2112 with a Standard In-Person Interview included eight participants who should have received the expanded version, as they had no CATI respondent. These eight participants were called back after the clinic to collect the additional residence history information that would have been collected in the Expanded IPI, in order to estimate their dose.

Table V.G-1. Summary of Standard and Expanded Interviews by Phase of Study

Version of	Pilot Study		Transition		Full Study		Total	
Questionnaire	No.	%	No.	%	No.	%	No.	%
Standard	750	70.6	427	64.3	935	54.4	2112	61.3
Expanded	313	29.4	237	35.7	785	45.6	1335	38.7
Total	1063		664		1720		3447	

G.7. Quality of In-Person Interview and Expanded In-Person Interview Data

Overall, the quality of the information obtained in the interview was judged by the Interviewers to be high. If the Interviewer assessed the quality of the data to be questionable or unreliable, then he or she recorded the reason for this determination and identified specific sections affected. Table V.G-2 shows the Interviewer's assessment of the reliability of the participant's responses to the standard and expanded versions of the questionnaire, as well as those Expanded In-Person Interviews used for dose estimation, i.e., excluding the interviews with insufficient residence history to calculate a dose estimate. Responses to the standard questionnaire were judged to be of high quality somewhat more frequently than those to the expanded questionnaire, but both versions were judged to provide high or generally reliable data in more than 95% of the interviews. Table V.G-3 shows the reasons for questionable or unreliable data. The most common reason was that the participant did not have a clear memory of the events in question. This reason was cited more often for the expanded version than the standard version. Approximately a quarter of the questionable or unreliable responses to both versions were due to an uncertain understanding of the questions by the respondent. All but a few participants were judged to have a very good or good level of cooperation (Table V.G-4).

Only 120 of the 3447 In-Person Interviews (both Standard and Expanded) were judged to have data of questionable, unreliable, or unknown reliability. Of these, 65 were used for dose estimation purposes. Note that in Tables V.G-2, V.G-3 and V.G-4, the third column, Expanded IPI Used for Dose Estimation, includes the eight participants mentioned above who had a Standard IPI at the clinic but should have received an Expanded IPI.

Table V.G-2. In-Person Interviewers' Assessments of Reliability of Responses

	Standard IPI		All Expanded IPI		Expanded IPI Used for Dose Estimation	
Reliability of Responses	No.	%	No.	%	No.	%
High	949	44.9	411	30.8	407	30.9
Generally reliable	1110	52.6	851	63.7	845	64.2
Questionable	49	2.3	61	4.6	59	4.5
Unreliable	1	0.0	2	0.1	2	0.2
Unknown	3	0.1	4	0.3	4	0.3
Expanded IPI done, data not used	--	--	6	0.4	--	
Total	2112	100	1335	100	1317	100

Table V.G-3. In-Person Interviewer's Assessments of Reasons for Questionable or Unreliable Information

	Standard IPI		All Expanded IPI		Expanded IPI Used for Dose Estimation	
Reason	No.	%	No.	%	No.	%
Unclear memory of events	16	0.8	34	2.5	34	2.6
Uncertain understanding of questions	15	0.7	15	1.1	15	1.1
Hurried responses	3	0.1	2	0.1	2	0.2
Other	16	0.8	12	0.9	10	0.8
Not applicable*	2062	97.6	1272	95.3	1256	95.4
Total	2112	100	1335	100	1317	100

* Reliability of Responses was High, Generally reliable, Unknown, or Expanded IPI done, data not used

Table V.G-4. Interviewers' Assessments of Respondent's Cooperation

	Standard IP		All Expanded IPI		Expanded IPI Used for Dose Estimation	
Respondent's Cooperation	No.	%	No.	%	No.	%
Very good	1703	80.6	1021	76.5	1011	76.8
Good	382	18.1	278	20.8	277	21.0
Fair	22	1.0	26	1.9	25	1.9
Poor	3	0.1	0	0.0	0	0.0
Not answered	2	0.1	4	0.3	4	0.3
Expanded IPI done, data not used	--	--	6	0.4	--	--
Total	2112	100	1335	100	1317	100

More than 90% of participants at least partially completed an Interview Preparation Worksheet prior to the interview.

The residence history in the Expanded IPI presented the most recall difficulty since participants were quite young at the time. Codes referred to as "fuzzy date codes" were assigned to each residence in the birth through 1957 section of the interview. The codes indicate the precision with which the participant was able to specify the date of a residence change (i.e. within two months, within three months, plus or minus 6 months, a year or more). This allowed coding of inexact responses to date questions to standard mm/yy codes. For example, responses such as "in the fall of 1947" and "1952 or 1953" would be coded as 10/47 and 1/53, respectively. These and other coding rules were contained in a written Coding Manual.

G.8. Conclusions

A complete In-Person Interview was obtained from all except six of the 3447 eligible study participants attending an HTDS clinic. These six non-evaluable participants were judged to have insufficient information in the residence history section of the expanded interview to calculate a dose estimate. The interview data were obtained easily and few modifications of the questionnaire were needed throughout the study. The Interviewers judged the responses to be Highly or Generally Reliable in over 95% of the interviews.

H. Medical Review and Final Diagnosis Determination

H.1. Background

H.1.a. Objectives of Medical Review and Final Diagnosis Determination

The objectives of the medical review and final diagnosis determination processes were: 1) to evaluate each participant's HTDS clinical thyroid evaluation results; 2) to communicate results of the clinical evaluation to participants in a timely manner and, with permission, to communicate the results to the participant's health care provider; 3) to assign the final diagnoses for each case according to the format developed (see Appendix 20) using all information available prior to and including the HTDS clinical evaluation.

H.1.b. Rationale

A large amount of information was collected for each study participant from the interview and clinical evaluation. This information included serum laboratory results; ultrasound exam, physical examination and for some participants, FNA results, thyroid nuclear scans and medical records. Members of the HTDS clinic team met in regular sessions to review the clinical information for study participants. The purposes of these reviews were to determine final diagnoses and to plan the letters and telephone calls for communicating the results to the participants and their health care providers. After all diagnostic information was assembled and the Medical Review was completed, Dr. Hamilton completed a Final Diagnosis Determination Form. This data form was used to record all of the final thyroid, parathyroid, or ultrasound outcomes of the HTDS diagnostic evaluation.

H.2. Medical Review and Final Diagnosis Determination

H.2.a. Medical Review Process

The results from the laboratory assays, cytology interpretations and radiologists' reviews of ultrasound tapes were received in the HTDS office within 5-6 days after the HTDS clinic. Results and review forms were assembled for each clinic participant for the weekly Medical Review session. The staff participating in the Medical Review included Drs. Hamilton and Griep, the Research Nurse, and the Field Operations Supervisor (FOS). During the review, the FOS completed tracking data forms for the Medical Review. Dr. Hamilton reviewed participants with no abnormal findings, while both Drs. Hamilton and Griep reviewed those with abnormalities.

All participants underwent a post-clinic Medical Review of the HTDS clinical evaluation results within two weeks of the clinic appointment. During the review session a letter to report the results of the evaluation was developed for each participant, plans were outlined for communicating abnormal results to participants, and a determination was made as to whether further diagnostic procedures or treatment should be recommended. If the participant did not report any history of thyroid disease during the In-Person Interview, a final diagnosis assessment was made and a Final Diagnosis Determination Form was completed as part of the Medical Review process. If a participant reported a past history of thyroid disease during the In-Person Interview, medical records were requested and the final diagnosis determination was deferred until after those records were obtained, abstracted, and reviewed.

H.2.b. Additional Tests

The first step in each participant's review was to determine whether any recommendations for further testing were necessary to confirm or rule out a diagnosis. Additional tests may have included: 1) thyroid nuclear scan (e.g. for Graves Disease or toxic thyroid nodule diagnosis); 2) repeat blood draw for additional tests such as parathyroid hormone in the case of elevated calcium; 3) repeat analysis of existing serum due to equivocal results; and 4) repeat thyroid FNA due to inadequate specimen.

An HTDS physician or the Research Nurse contacted participants needing additional tests by telephone to discuss the abnormal results, to answer questions, and to recommend the appropriate follow-up procedures. If consent had been given to contact the participant's health care provider, that person was also contacted to discuss the recommendations. Following this initial contact, the Research Nurse re-contacted each of these participants on a regular basis to determine if the recommendations had been followed, and to obtain consent for receiving results reports. If consent was obtained, medical records, cytology and pathology slides and reports were requested and reviewed in the same manner as historical records and slides. The tracking system was used to track the progress of recommendations for further procedures and the acquisition of outcome information.

Nuclear scans were arranged by the Research Nurse to be done at a medical facility most convenient for the participant. For additional blood tests, the participant's blood was collected at a local health care provider's office or the nearest medical laboratory and shipped to Pacific Medical Center Clinical Laboratory in Seattle for processing. The Research Nurse handled all arrangements for follow-up tests, shipping of specimens and payment of services.

H.2.c. Communication of Medical Review Results to Participants and Their Health Care Providers

After evaluating each participant's clinical information, the physicians drafted a letter to the participant outlining the results of the evaluation. The Data Control Technicians printed the letters and attached the laboratory results pages and appropriate fact sheets for each participant. If the participant's results were all normal, the results letter was mailed out immediately. The Research Nurse entered information about all cases with abnormal findings into a follow-up system for contacting by telephone and further follow-up as indicated. The results letters for the participants with abnormal findings were mailed after telephone contact by the Research Nurse. All participants received their results within 3-4 weeks after their clinic appointment.

Letters were also sent to each participant's health care provider if the participant indicated this was to be done and supplied the provider's name and address. The letters to health care providers included recommendations for follow-up monitoring and tests. The health care providers were also sent copies of the results letter and the fact sheets sent to the participant.

All participants who had an FNA recommended by the HTDS physicians were called by the Research Nurse or Dr. Griep on the day of the Medical Review. If the participant gave consent, the health care provider was also contacted by telephone to discuss biopsy results and to answer any questions. On the rare occasion when a repeat FNA was recommended, participants were called by Dr. Griep to discuss the results of the procedure.

If additional tests were recommended after the Medical Review, the results of these tests were reviewed at the next Medical Review session and a second results letter was mailed to the participant and his/her health care provider, describing the results of the follow-up tests.

H.2.d. Fact Sheets

Fact Sheets on various topics related to the HTDS clinical evaluation and results were developed and distributed to provide information to study participants. One fact sheet described the purpose and explained the results of the blood tests conducted as part of the clinical evaluation. A Physician Referral Resources handout was developed to help participants locate a health care provider for follow-up on conditions identified in the HTDS evaluation. A fact sheet explaining autoimmune thyroiditis, and a series of fact sheets about nonpalpable ultrasound-detected abnormalities of the thyroid were provided only to participants with these results. An Ultrasound Follow-up fact sheet was distributed to explain the follow-up program for persons with ultrasound-detected abnormalities of the thyroid. Following discontinuation of the Ultrasound Follow-Up Program in April 1995 (see section V.F.5 above), the Ultrasound Follow-up fact sheet was discontinued and an additional paragraph of information was added to the results letter to discuss the significance of nonpalpable ultrasound detected thyroid abnormalities. Later, a new fact sheet describing what was known about nonpalpable ultrasound-detected thyroid abnormalities was added to provide additional information for participants (and their health care providers) who had no other thyroid disease identified.

H.2.e. Final Diagnosis Determination

Findings from the HTDS clinical evaluation, and in some cases, historical medical records identified during the interview process (see section V.G), were reviewed to determine the participant's final diagnoses. The final diagnoses included information about thyroid and parathyroid outcomes (including basis for diagnosis) and ultrasound findings. Diagnoses for cases with no indication of thyroid abnormalities were made by Dr. Hamilton. Diagnoses for cases with any indication of an abnormality were made by consensus of Drs. Hamilton and Griep. The final diagnosis data were recorded on the Final Diagnosis Determination Form (see section IV.C above). Final diagnoses or disease outcomes were further defined by variables to indicate the quality and source of documentation on which the diagnosis was based.

Final diagnosis determinations were made based on the following information: 1) HTDS blood test results; 2) HTDS ultrasound results; 3) HTDS examination results; 4) previous thyroid disease or treatment with thyroid medication reported in the HTDS In-Person Interview or CATI; 5) current use of thyroid prescription medication reported in the HTDS In-Person Interview; 6) HTDS FNA results, if any; 7) HTDS-recommended diagnostic or surgical procedures, if any; and 8) historical medical records obtained by HTDS, if any.

The Final Diagnosis Determination Form underwent minor revisions during the first two years of the study. In July 1995 the following three significant changes were made: 1) the two outcomes for multinodular goiter based on being on thyroid medication or not, were consolidated into one category designated multinodular thyroid gland; 2) Graves Disease was added as a separate diagnostic outcome; and 3) the "basis for diagnosis" and "Histologic/Cytologic Type" sections for each diagnosis were expanded and standardized.

After these revisions, it was necessary to review and revise the Final Diagnosis Determination Forms for approximately 1376 participants whose diagnoses had been assigned on earlier versions of the form. Six staff members reviewed the original Final Diagnosis Determination Forms and transferred data to the new forms. Laboratory results from the clinical evaluation were reviewed and ultrasound findings documented for each case. If no diagnoses had been indicated (i.e., findings were normal) the revised form was considered complete after verification by a second staff person. Dr. Hamilton thoroughly reviewed the diagnosis determination for cases with findings varying from those originally documented and cases where at least one diagnosis had previously been identified.

H.2.f. Dating of Diagnoses

To perform analyses accounting for the potential effects of the Nevada Test Site exposure, diagnoses made before 1957 had to be distinguished from later diagnoses. Therefore a date was assigned for every diagnosis recorded on the Final Diagnosis Determination Form, corresponding to the date or age of that diagnosis. If there were medical records, or a prior mention of thyroid disease during the CATI or In-Person Interview, the subject's chart was reviewed to determine the date or age of diagnosis. Otherwise, if the diagnosis was based on findings at or as a result of the HTDS examination, the clinic appointment date was assigned as the date of diagnosis. When dates/ages were not specific, the midpoint of the range was assigned as the date/age of diagnosis.

H.3. Outcome and Results

H.3.a. Number of Cases Reviewed and Follow-up Procedures Recommended

The total of 3447 eligible participants underwent medical review. For 79.3% of these participants, the Final Diagnosis Determination Form was completed at the time of their Clinic Medical Review. The remaining 20.7% had either requests for historical medical records or post-clinic recommendations for further diagnostic procedures. For these cases, the Final Diagnosis Determination Form was completed after all the additional results or records had been received.

A total of 259 participants had FNA procedures performed at the HTDS clinic or on the recommendation of the HTDS physicians after the Medical Review. Of these 259 participants, the HTDS physicians recommended that 47 participants have further biopsy or surgical procedures to rule out a diagnosis of thyroid neoplasm, or were recommended to undergo close follow-up by their health care provider to monitor progression of a thyroid disorder. Another fifteen were followed for further diagnostic tests such as blood redraws or nuclear scans. In addition, 29 participants with thyroid nodules or suppressed TSH were recommended to undergo thyroid nuclear scan. Twenty participants had an abnormal calcium level and were recommended to have additional blood drawn and analyzed for parathyroid hormone (PTH) studies to confirm or rule out a diagnosis of hyperparathyroidism. Thirty participants were requested to have additional blood drawn due to abnormal or borderline thyroid function.

H.3.b. Conclusions

The Medical Review and final diagnosis determination processes were conducted efficiently and all participants received a thorough evaluation of the clinical results to determine the presence or absence of thyroid disease. Study participants were provided with their results in a timely and considerate manner and they were provided with recommendations for follow-up if a condition was identified.

I. Historical and Post-Clinic Medical Records and Specimens

I.1. Background

I.1.a. Objectives of Obtaining Medical Records

The primary objectives of the medical record component were to: 1) document thyroid problems reported by study participants and CATI respondents; 2) obtain any cytological or histological specimens from previous thyroid biopsies or surgeries for review by the study's pathologist; and 3) obtain the results (including histological specimens) of any further diagnostic or surgical procedures recommended by HTDS as a result of findings at the HTDS clinic. A secondary objective of the medical record component was to obtain cause of death information on all cohort members located deceased, and assign cause of death codes according to a standardized rubric.

I.2. Process and Procedures Used

I.2.a. Historical Medical Records

Information was obtained from both the participant and the CATI respondent for the purpose of obtaining historical medical records. During the CATI, respondents were asked to provide the names (and addresses, if known) of any physician who saw the participant for diagnosis or treatment of thyroid disease. Prior to the clinic appointment, the participant was sent a work sheet on which to list the names and addresses of their current physician and any previous physicians seen for diagnosis or treatment of thyroid disease. At the time of the In-Person Interview, the participant was asked to provide the names and addresses of physicians or institutions where they had been diagnosed or treated for thyroid or parathyroid disease, and to sign a consent form for the release of information from each of these providers. Information from the CATI was provided to the In-Person Interviewer in the clinic packet so that consent for any records identified only by the CATI respondent could also be obtained from the participant during the clinic visit.

All completed consent forms were returned to the office, where a Data Technician reviewed them for completeness. If the provider's address was unknown to the participant, attempts were made to locate a current address so the consent could be delivered. A letter requesting the pertinent records was generated to accompany each consent form. Copies of the consent forms were filed in the participant's medical record. If a current address for a provider could not be obtained, the original consent signed by the participant was filed in the record with a notation that the provider could not be located. A log of medical record requests was kept so that requests could be followed and further action taken if records were not received. Information regarding the request of medical records was updated in the study Tracking System.

Once records were received, they were given to the Medical Records Abstractor for review, organization, and abstracting of laboratory values. If specific records were found to be missing, the Abstractor relayed this information to the Data Technician, who re-contacted the provider for the information. Once the record was deemed complete for HTDS purposes, the abstract and records were filed in the HTDS medical record, and the study Tracking System was updated to indicate the case was ready for Medical Review.

During the Medical Review sessions (for cases with medical records), the Medical Records Supervisor and Dr. Hamilton reviewed each case and Dr. Hamilton assigned the proper final diagnoses.

I.2.b. Post-Clinic Medical Records

For those participants where HTDS physicians recommended further work-up or treatment based on the HTDS clinic findings, medical records documenting these procedures were also requested. At the time the recommendation was made, the Research Nurse asked the participant to give consent for the HTDS to obtain these records. Once signed consent forms were received (either at the clinic or through the mail, if the recommendation was made based on the results of clinic cytology or blood tests), they were handled in the same manner as those for historical medical records. Tracking of requests, however, was handled by the Research Nurse, who kept in contact with the participants throughout any further evaluation, to ensure that recommendations were adequately carried out, or that the participant fully understood the ramifications to their health if they chose not to do so. Once records were received, this information was entered into the study tracking system, and the records were flagged as ready for medical review.

I.2.c. Blinding the Reviewer to Radiation Exposure

References to areas in which the participant had lived were blocked out of any records prior to review by Dr. Hamilton so no inference of radiation dose could be made based on past residences. This blinding of radiation exposure was accomplished in the following manner. When a record was found to state that the participant was a "downwinder," had lived in the Hanford area (or an area away from Hanford), or had been exposed to radiation from Hanford, a copy of those records was made. The original was filed in a section of the HTDS medical record marked "Unused Information," and stapled with a cover sheet so it would not be read inadvertently. The exposure information on the copy was then deleted, and the blinded copy used in the Medical Review process by Dr. Hamilton. The participant's name was recorded in a logbook along with information identifying the records that mentioned exposure status.

A similar procedure was used for records that indicated the participant had undergone radiation therapy for malignancies other than thyroid. In cases where the participant or respondent reported thyroid disease in the participant, only radiation therapy for disseminated malignancies such as leukemia was blinded. For participants without any report of thyroid disease, only radiation therapy to the upper body was blinded, including any radiation for disseminated malignancies such as leukemia. Again, details on the blinded records were recorded in a logbook.

Upon instituting these procedures, it immediately became apparent that by only blocking out such references to radiation exposure, Dr. Hamilton might infer that all cases with sections censored had references to radiation exposure. Thus it was decided that some random blinding of records would need to be performed as well. For this reason, every seventh case with medical records was selected for random blinding. The information blinded in these cases was always completely unrelated to radiation exposure status or radiation therapy. For example, references to previous gallbladder surgery might be censored in one case; while in another, documentation of a motor vehicle accident might be censored. Again, the original records were placed in the "Unused Information" section, with the blinded copies used for medical record review.

I.2.d. Cause of Death Coding

For each potential participant located deceased, the death certificate or informant information was used to complete a Cause of Death Form (Appendix 21). In addition, the primary cause of death was coded using the ICD9-CM system. For those whose date of death preceded the use of the ICD9-CM system, the primary cause of death was also back-coded using the system in use at the time of death (Table V.I-1). See section V.B.4.d.3 above, for detail on the success in obtaining death certificates.

Table V.I-1. Systems Used for Cause of Death Coding

Coding System	Date Published	Dates of Use
International List of Causes of Death	1939	1940-1948
International Select Causes of Death	1948	1949-1955
International Classification of Diseases, Sixth Revision	1955	1956-1961
International Classification of Diseases, Seventh Revision	1962	1962-1967
International Classification of Diseases, Eighth Revision	1968	1968-1978
International Classification of Diseases, Ninth Revision	1979	1979-1997

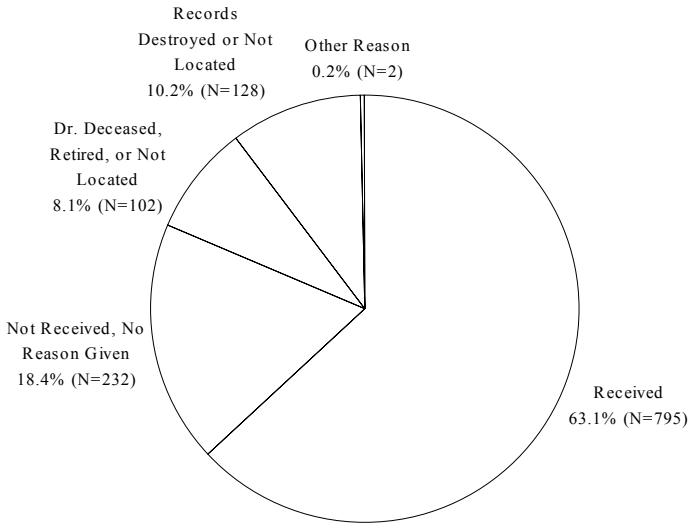
I.3. Outcome and Results

I.3.a. Historical Records

Reports of historical medical records were obtained for 694 participants, with a total of 1259 consent forms completed to obtain medical records from different providers. While the vast majority of reports were made during the In-Person Interview, CATIs yielded 30 of these reports.

Of the 1259 Medical Record Consents obtained, a total of 795 (63.1%) separate medical records were received by the HTDS. No records were received for 464 requests (36.9%). Figure V.I-1 shows the reasons for non-receipt of records. In 102 (8.1%) cases, records could not be requested because the physician was deceased, retired or a current address could not be identified. For 128 (10.2%) requests, records were unavailable due to the destruction of records, the inability of the provider to identify the patient, or an inability to locate the records. In 232 (18.4%) cases, records were not received after several contacts, without explanation as to why they were not available.

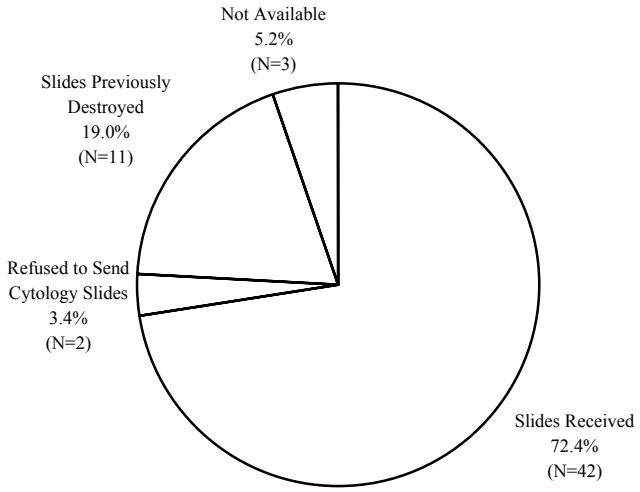
Figure V.I-1. Outcome of Historical Medical Record Requests (N = 1259)



I.3.b. Historical Pathology and Cytology Slides

Of the 694 participants identifying historical medical records to be requested, pathology or cytology slides were requested for 52 (7.5%). In a few cases, more than one set of slides was requested, for a total of 58 separate requests. A total of 42 sets of historical pathology or cytology slides were received for 42 participants (80.8% of those for whom slides were requested). Of the requests not resulting in receipt of slides, 11 were related to procedures performed prior to 1985 and the slides had been discarded, and three were not available. In the other two cases, cytology slides were not provided, as there would be no additional slides on file if they were lost or not returned. In these two cases, copies of the cytology reports were provided, and it was determined that the cytology appeared to be from the same nodules on which FNA had been performed at the HTDS clinics. Figure V.I-2 shows the success in obtaining historical cytology and pathology slides for review.

Figure V.I-2. Success Obtaining Historical Slides (N=58)



I.3.c. Post-Clinic Medical Records and Slides

Medical records documenting further diagnostic studies recommended as a result of the HTDS clinic findings were requested for 35 participants, with a total of 72 separate requests. All but one of these records was obtained, with at least one record obtained for each of the 35 participants. Thirty-three of these participants also had histology or cytology slides requested, for a total of 35 separate requests. All thirty-five of these specimens were obtained.

I.3.d. Cause of Death Coding

Cause of death was coded for 527 potential participants who were located deceased and 16 potential participants who were located alive but died prior to participation in the HTDS. Death certificates were received for 504 of the 543 total deceased potential participants.

In no case was thyroid disease listed as a primary or contributing cause of death on any death certificate obtained. In one case, the family member informant stated that the potential participant's cause of death may have been thyroid disease. However, the death certificate was obtained for this potential participant and the primary cause found to be malignant lymphoma.

I.4. Potential Impact of Medical Records and Slides That Were Not Obtained

One potential concern is that diagnoses of disease outcomes might be missed if requested medical records or slides could not be obtained: none or only part of the requested records or slides were received for 199 (29%) and 160 (23%), respectively, of the 694 participants for whom such requests were made. However, even if a medical record or slide could not be obtained, the likelihood of a missed diagnosis was generally low because in most such situations the HTDS evaluation provided a definitive assessment of whether the diagnosis for which the medical record or slide was sought was confirmed or not confirmed. For example, if a participant or CATI respondent reported a diagnosis of a thyroid nodule 30 years ago, that diagnosis would almost certainly be confirmed by HTDS physicians based on current physical exam and ultrasound scans. An exception would be for a participant reporting thyroid cancer, who then had thyroid surgery, and then had missing medical records. However, this occurred in only one individual**.**

To further clarify this issue, all of the diagnoses for participants with at least one missing medical record or slide were reviewed to determine which did not have a confirmed HTDS diagnosis and therefore might have been missed based on not receiving that medical record or slide. Of the 556 diagnoses for the 359 participants with at least one requested medical record or slide not obtained, 318 (57%) were confirmed by the HTDS evaluation. Of the remaining 238 diagnoses, 109 were diagnoses of hypothyroidism for which at least one requested medical record or slide was not obtained. Each of these 109 participants had normal thyroid function during the HTDS evaluation, thereby eliminating the possibility of permanent hypothyroidism. Of the remaining 129 diagnoses, 54 (16 with hyperthyroidism, 28 with simple goiter or multinodular gland, and 10 with "other" thyroid disease) were reported by or for participants who had completely normal HTDS thyroid evaluations. In addition, there were four miscellaneous reports of thyroid disease for which the HTDS evaluation was normal. For these 58 diagnoses the normal HTDS evaluation eliminates the possibility that these participants had permanent thyroid disease in these categories. Consequently, there were 71 diagnoses (13% of the 556 diagnoses) for which a missing record might have contributed to a diagnosis that was not confirmed by the HTDS evaluation. It must be emphasized however, that given the completeness of the HTDS clinical evaluation, this figure of 13% is likely an upper bound for the possibility of missed diagnoses related to any missing medical records.

Finally, it should be noted that the analysis of each thyroid disease outcome and of hyperparathyroidism included estimation of the dose-response not only for the definitive diagnoses based on HTDS evaluation or medical records with documentation supporting the diagnosis, but also for those based on less definitive criteria. In particular, these alternative analyses included diagnoses based on CATI respondent or participant reports only, which by definition had no confirmation from either the HTDS examination or from any medical records.

I.5. Conclusions

Attempting to obtain medical records and slides from as long ago as fifty years prior was expected to be one of the most difficult aspects of the HTDS. Many medical records are destroyed after only seven to ten years; physicians retire, sell their practices, or die, leaving little hope of locating historical records and slides. While no estimates of success in locating such records or slides was made in the HTDS Protocol, it was generally felt that records and slides would be obtained in no more than 50-60% of cases. While the experience of the HTDS was only slightly better at 63.6% of consents resulting in records or slides, it should be noted that of 694 participants with historical records or slides requested, 495 (71.3%) of participants identifying one or more records or slides had at least one record or slide retrieved. More recent

records and slides were retrieved more easily and, in many cases, these referred to earlier diagnoses for which the original records or slides could not be located. This enabled the study to confirm historical diagnoses in a greater percent of cases, despite the lack of older records.

J. Data Management

J.1. Objectives of Data Management

The primary data management objective was to establish procedures that would be used to develop and maintain the study databases, and the procedures that would be used to ensure data quality. These procedures included manual review (editing) of data recorded on paper forms, duplicate entry for all data forms, validity checks encoded in the data entry programs, and consistency check programs run on the data after entry.

The second data management objective was to define procedures to maintain the security and confidentiality of the data. This included data in computerized form, through the use of passwords and control of limited access to directories and data files, as well as hard copies of data, i.e., paper records, which were stored securely in locked files in locked offices or in a file room which had limited access via keycard.

The data collected for this study were classified into six main categories for purposes of data management:

1. Tracking system
2. CATI
3. In-Person Interview
4. Clinic Data Forms, Final Diagnosis Determination Forms, Refusal Questionnaires, Cause of Death Forms and Dating of Diagnoses
5. ICD9 Coding of Cause of Death
6. Problems Forms

The general principles guiding the development of data management procedures were the same for all six categories of data. These general principles included the use where appropriate of manual reviews (editing) of data originating on paper forms, duplicate data entry, automatic validity checks at data entry, additional computerized checks of data validity, frequent backups of computerized databases, password control of access to databases and of authority to update databases, and restricted access to physical repositories of data and specimens. The specific implementation of these general principles varied according to the nature of the data in each category, as described below.

At the beginning of the study, transfer of data between computers or from computers to back-up storage media was accomplished by means of removable media such as diskettes or tapes. Subsequently all of the study computers were connected to a local area network (LAN), and thereafter transfers and back up of data were managed through the LAN. The descriptions that follow describe the procedures that were adopted following availability of the LAN.

J.2. Data Management Procedures

J.2.a. General Procedures

Data were entered on personal computers which were linked via a LAN. Some data entry programs were stored on this network, in a directory with limited access. The LAN was backed up each business day by the network administrator.

To maintain confidentiality of the data, multiple levels of security were employed. First, the HTDS staff worked in a secure building with access to the floor limited to those with a security key card. The actual computers used for data entry were kept in locked offices and only authorized study personnel

had access to them. Most of the study computers were linked via a Novell local area network (LAN) and access to the data entry programs located on the local area network was limited via user names. The data files stored on the network were backed up daily Monday through Friday by the network administrator. All other data entry files were backed up daily when in use on floppy diskettes.

Apassword program was installed on all study computers. The password was changed whenever an employee left the HTDS or periodically during times when there were no staffing changes. As an additional security measure when an employee left the study, the network administrator revoked access to the local area network. A final security measure utilized was to keep all completed study forms in locked files or in a fileroom with limited key card access.

Preprinted labels with the participant's unique identification (ID) number were attached to each of the participant's forms to prevent transcription errors.

J.2.b. Tracking System

The tracking system contained data regarding the progress of participants through the study. These data were used by study staff to ensure the prompt and complete progress of potential participants through the various components of study participation: recruitment, identification of a CATI respondent and completion of the CATI, clinic scheduling, completion of clinic activities and recommended follow-up procedures, medical records requests, and completion of medical review and participant contacting. The tracking system database was written using dBase IV software as a menu driven program consisting of eight databases linked by the participant's ID number and last name. Table V.J-1 gives a brief description of the eight databases in the Tracking System:

Table V.J-1. Tracking System Databases

Database	Description
Overall	Summary of each potential participant's status.
Tracing	Tracing outcome, vital status, and death information.
Participation	Agreement/refusal/withdrawal information.
Dosimetry	Information regarding the CATI respondent and agreement and completion status of the CATI. Contains more than one record for some participants.
Clinic	Clinic appointment information and completion of individual clinic items.
Miscellaneous	Information regarding blood re-analyses, nuclear scans and repeat FNAs. Contains more than one record for some participants.
Medical Records	Information regarding medical records requests. Contains one record per medical record request, and thus has more than one record for some participants.
Participants	Participant's name and identification number.

Data entry programs for the tracking system were written to be user friendly, with all appropriate instructions on the screens as needed to enhance ease and accuracy of use.

Entry of data into the tracking system was not done in duplicate. Duplicate entry turned out to be impractical, and was deemed unnecessary in part because these were not outcome data. However, a program was run on these data periodically to check for invalid or inconsistent data.

Information about located potential participants was initially entered into the Tracking System on a weekly basis by a Data Technician, using data received from the Tracing staff in the Tri-Cities. Once the data for the week were entered, an electronic mail message was sent to key HTDS staff alerting them that the Tracking system had been updated, and noting any special circumstances for the new potential participants, who were identified only by ID number. As the various steps involved in contacting, recruiting, identifying a CATI respondent and completing the CATI, and clinic scheduling of a potential participant were accomplished, study staff created and updated records in the appropriate Tracking System data bases.

Following each clinic, the list of participants attending, and the steps of the clinic process completed by each participant were entered into the system by the Data Technicians. For participants who identified historical medical records, or for whom additional post-clinic studies were recommended, records were created in the Tracking System to track the request and receipt of these data, and to flag cases as ready for Medical Review.

J.2.c. CATI

A Computer Assisted Telephone Interview (CATI) was used for data entry of the dosimetry questionnaire on a real-time basis during the interview. The data entry program for this interview was written by the study programmer using the INGRES database software. Since the responses received during the CATI were entered directly into the database, they were not verified by duplicate data entry. One section that was handled somewhat differently was the residence history section, which was sent to the CATI respondent prior to the telephone interview. In most cases, the residence history was returned and data entered prior to the actual telephone interview, and then reviewed in the course of the interview.

The CATI database was programmed with automatic range checks, as well as skip patterns, where appropriate. (In the process of creating the scenario files used for dose calculations, additional consistency checks were run on the CATI data.)

At the end of each day during which they performed CATIs, each Interviewer copied the CATI database from his or her data entry computer to the network. This computer was equipped with a tape backup unit and software capable of making unattended backups at pre-selected times. See section VI Dose Determination, for more details regarding the CATI data.

J.2.d. Clinic In-Person Interview

The In-Person Interviews were completed and reviewed (edited) by the Interviewers at each clinic. After each clinic, the Field Operations Coordinator again reviewed (over-edited) the interviews and coded all items except the grid locations in the residence history section. A data technician, using a map provided by Battelle PNL for this purpose, coded locations of residences from the Residence History. Once the locations were coded, the interviews were data entered and verified in an INGRES database by a data technician on a separate personal computer. The verification database included programmed range checks and appropriate skip patterns. Upon completion of verification, the interview data were copied onto the local area network. These data were then converted into SAS® databases using the software package DBMSCOPY. Programs were written in SAS® to perform additional consistency and edit checks.

There were two types of the in-person interviews: the Standard In-Person Interview (designed for participants for whom a CATI had been completed) and the Expanded In-Person Interview (designed for

participants for whom a CATI had not been completed). A total of six versions of the Standard and Expanded In-Person Interviews were used during the study. The differences between versions consisted primarily of minor wording changes, and when revisions to the questionnaire were made, the appropriate changes were made to the data entry program. After the Pilot Study, however, there was a major revision to the way the residence history questions were asked. See section V.G above for details of the revisions made to the questionnaire. Due to the major revision after the Pilot Study, new versions of the In-Person Interview database and data entry programs were created. When the data technician copied the files onto the local area network after this revision was in place, two new directories were used, one for the Standard and one for the Expanded version of the questionnaire.

J.2.e. Clinic Data Forms, Final Diagnosis Determination Form, Refusal Questionnaire, Cause of Death Form and Dating of Diagnoses

Eight data forms were completed as part of the clinic component of the study. These forms included: 1) Clinic Flow Sheet, 2) Thyroid Exam Form, 3) Consensus Exam Form, 4) Post Ultrasound Consensus Exam Form, 5) Ultrasound Form, 6) FNA Form (as needed), 7) Blood Test Results Form, and 8) Final Diagnosis Determination Form. The Refusal Questionnaire, completed for those who refused participation but agreed to answer demographic questions, the Cause of Death Form, and Dating of Diagnoses were processed in the same manner as the clinic component data forms and are therefore included in the following descriptions.

The data forms were entered and verified in SPSS data entry files. The data entry programs contained range checks as well as skip and fill rules. A data entry manual was written for each data form, and included step-by-step instructions for entering the data. These manuals also contained detailed information regarding the variable names, types, length, description, and valid codes, and outlined the skip patterns. At the back of each manual was a Decision Log, where the Data Entry Operator identified cases that required a decision regarding how to enter the data. The Data Supervisor or the Statistical Research Associate reviewed all entries in the decision log. The data entry manuals were stored in the Data Supervisor's office.

At the end of each month, the SPSS data entry files were converted to SAS® data files via DBMSCOPY software that was available on the local area network. The converted files were then checked for duplicate records and unverified records, and any problems found were corrected. These files were then appended to the master files, which contained all previously cleaned data. As these files were appended, they were also compared to the master files for duplicates, and any found were deleted. The master files were stored in a directory with limited access, so that only the Data Supervisor and Statistical Research Associate could make changes to the files.

After the files were appended to the master files, a series of SAS® programs were run to check for any inconsistencies in the data files, such as skip patterns, invalid values, etc. When an inconsistency was found, the original clinic data forms were visually checked and a SAS® program was written which made the appropriate corrections. Hard copies of all such programs were stored in a locked file cabinet in the Statistical Research Associate's office.

Most of the clinic forms underwent minor revisions during the course of the study, and when necessary, the data entry programs were revised accordingly after careful investigation to ensure consistency of pre- and post-revision data. More significant modifications are described below:

Clinic Flow Sheets: The clinic flow sheet was used to track the progress of participants through the clinic, and did not contain substantive information regarding participant characteristics, radiation exposure, or outcomes. Data entry of the clinic flow sheet was discontinued after the 5/20/95 clinic.
FNA Form: The FNA form was not originally planned to be data entered. However, it was determined that it would be of interest to know how many FNA procedures were performed by each physician and

consequently the FNA forms were data entered. In addition, an FNA form abstract was developed and used for the few FNA procedures that were performed at the request of HTDS but by an outside physician. This form indicated the number of nodules aspirated by the outside physician. The ID number, date of FNA and number of nodules aspirated was entered, in order to count how many nodules were aspirated.

Blood Results Form: The blood test results were reported from the laboratory on their own standard form. The database and data entry programs were revised when necessary to accommodate changes in the laboratory's form or in the assays they performed. Additional data were entered by study staff to further characterize the laboratory results, including identifiers of reassays (initial result, or reassay of original specimen), TSH assay type (RIA, EIA-1, or EIA-2), and PTH and calcium assay types (IRMA or chemiluminescence). See section V.F.3 Clinics, for further information on the laboratory assays used.

Final Diagnosis Determination Form: This Form underwent significant revisions during the course of the study, and the last version was adopted on July 28, 1995. To ensure consistency of the data from this key form, a copy of the final version was completed for all participants, including those whose medical review results had been recorded on an earlier version of the form. Please refer to section V.H.2.e above for a description of the changes made to the Final Diagnosis Determination Form.

J.2.f. ICD9 Cause of Death Coding

The ICD9 Cause of Death Coding was entered into an Excel spreadsheet by the staff member who performed the coding and then verified and resolved any discrepancies. This spreadsheet was then converted into a SAS® datafile using the software package DBMSCOPY.

J.2.g. Problems Forms

Throughout the study, staff were encouraged to bring any procedural problems to the attention of the supervisors and the Study Management Team. To formalize this process, a Problems Form was created. This form was completed by the person(s) who identified the problem and given to the Administrative Coordinator, who was responsible for bringing it to the attention of the appropriate people to be involved in determining and carrying out the resolution. The Administrative Coordinator was responsible for tracking progress toward resolution of items on Problems Forms, and prepared a weekly summary of the status of the outstanding forms. This summary was reviewed by the supervisors at weekly meetings and was also given to the Study Management Team for review. As resolutions to each problem were decided upon, these were logged by the Administrative Coordinator and the resolution of the problem was recorded on the Problems Form.

J.3. Outcome and Results

J.3.a. Tracking System

Table V.J-2 displays the number of records in each of the tracking system databases.

Table V.J-2. Number of Records in Each Tracking System Database

Database	No. of Records
Overall	4346
Tracing	4883
Participation	4348
Dosimetry	4447
Clinic	4346
Miscellaneous	69
Medical records	1443
Participants	4385

J.3.b. CATI

A total of 2133 participants who attended the HTDS clinic had a CATI. Of these 2133 participants, 29 also had an Expanded In-Person Interview. An additional 135 potential participants had a CATI, but withdrew from the study before attending an HTDS clinic.

J.3.c. Clinic In-Person Interview

A total of 2112 participants had a Standard In-Person Interview at the clinic. In addition, 1335 participants had an Expanded In-Person Interview at the clinic. One of these participants had help from his/her father as well as a caregiver during the Expanded In-Person Interview.

J.3.d. Clinic Data Forms, Refusal Questionnaires, Cause of Death Form, and Dating of Diagnoses

Table V.J-3 indicates the number of each of the clinic data forms, Refusal Questionnaires, Cause of Death Forms and Dating of Diagnoses that were data entered and the number of people with at least one of these forms entered for the 3447 eligible participants who attended a clinic.

Table V.J-3. Numbers of Records in the Clinic Database, Refusal Questionnaires, Cause of Death Forms, and Dating of Diagnoses

	Data Entered and Verified	
	No. of Forms	No. of Participants
Clinic Data Forms		
Clinic Flow Sheet	1192	1192
Initial Blood Test Results	3439	3439
Thyroid Examination Form	6899	3447
Consensus Examination Form	3447	3447
Ultrasound Form		
▪ Ultrasonographer	3447	3447
▪ Radiologist	3447	3447
▪ Ultrasonographer QC	103	103
▪ Radiologist QC	343	329
Post Ultrasound Consensus Exam Form	3448*	3447
FNA Form	263	259
Final Diagnosis Determination Form	3447	3447
Additional Blood Test Results		
Thyroid Function Redraws	27	27
Calcium Function Redraws	20	20
Reanalysis of Clinic Panel	37	37
Reanalysis of Thyroid Redraw	0	0
Reanalysis of Calcium Redraw	1	1
Refusal Questionnaire	365	365
Cause of Death Form	543	543
Dating of Diagnoses	1258**	667

* Includes one participant with a second Post-Ultrasound Consensus Exam Form following review by the radiologist, which indicated a new nodule.

** Number represents total number of diagnoses assigned a date prior to the clinic appointment date.

J.3.e. Problems Forms

A total of 147 Problems Forms were completed during the study. Problems ranged from clerical scheduling issues to final diagnosis determination. Each problem was reviewed and possible solutions discussed by the appropriate staff. All problems on the forms were ultimately resolved.

K. Data Quality Control

In addition to the data management plans and procedures outlined in the previous section, additional steps were taken after data collection to ensure a high degree of data quality. These efforts included more extensive examination of the In-Person Questionnaire data, the CATI data, scenario file Construction, dose estimation, and computer programming. More detailed descriptions of these efforts follow.

K.1. In-Person Interview Questionnaire Data

Data from the In-Person Interviews were entered into four INGRES databases, one each for the original and the revised versions of the Standard and Expanded In-Person Interviews. Each database consisted of approximately 25 tables. The frequency distribution of each variable was examined to check for invalid codes or values. Data from questions involved in skip patterns, i.e., questions that that might or not be asked, depending on the response to another question, were also reviewed. If any problems were found, such as invalid codes or inconsistent skip patterns, the participant's questionnaire was reviewed and the correct code was entered into the database. The background table, which includes the participant's ID, date of birth, and date of interview, was compared to the tracking system database to assure that each study participant who completed an In-Person Interview was present in the background table. Participant ID number then linked all of the tables. The background table was used as a reference to confirm that participants had information in each table. Once all of the within-table and between-table checks were completed, the four databases were compared to check for duplicate ID numbers between databases.

The most frequent problems found as a result of reviewing the In-Person Interview databases included the following: (1) residence histories having multiple addresses with the same move-in and/or move-out dates, and (2) medical histories reporting an age at first medical procedure less than 15 years of age in the Standard In-Person Interview, which only asked about procedures after the age of 15. The questionnaires of participants with either of these data problems were reviewed and the database revised as needed. In addition, several database tables contained exactly duplicated records due to a programming error in the INGRES database structure. Finally, many extraneous records had been created when previous attempts to correct records in the databases resulted in the creation of new records, rather than overwriting of records. These duplicate and extraneous records were identified and deleted.

K.2. CATI Data

The CATI data were stored in INGRES databases, each consisting of approximately 65 tables. Quality control began within each table by reviewing the frequency distribution of each variable. If any invalid codes were found, the audiotapes of the pertinent interviews were reviewed and the database was corrected accordingly. Data from questions involved in skip patterns were checked for consistency. Once these within-table checks were completed, the log table, which contains the participant's ID number, date of birth and date of interview, was compared against the tracking system to assure that all participants for whom a CATI had been completed appeared in the CATI databases. This table was then used as a reference in the between-table checks. When possible errors or inconsistencies were detected, the audiotape of the pertinent CATI was reviewed and the CATI database was updated as needed. Once all of the within-table and between-table checks were completed, the two databases were merged to search for any duplicate participant ID numbers.

After the two CATI databases were merged into one database, a series of more in-depth computerized consistency checks were conducted. These included the identification of definite inconsistencies as well as suspicious data that weren't necessarily inconsistent. All definite and possible

inconsistencies were investigated by reviewing the audiotapes of the CATIs, and correcting the CATI database as needed. Below is a more detailed description of the additional checks that were performed.

- A comparison was made between the table indicating the types of milk consumed at each residence to the tables containing the actual milk consumption values, to ensure that only the types of milk specified for a particular residence were reported as having been consumed at that residence. Note that data from the table indicating milk types is not used in the dose calculation, since this information is implicit in the information about quantities consumed. It was known that CATI respondents sometimes changed their responses during the interview, and it was believed that when they did, the tables of actual consumption values of the various types of milk would be the most accurate. Many of the discrepancies detected resulted in revision in the table of milk types. However, this comparison also identified occurrences of a particular data entry error that occurred when the Interview sometimes neglected to zero out of the consumption of a given type of milk, when the discontinuation of that type of milk occurred at the time of a residence change. For example, a participant may have consumed raw cow's milk only at his or her first residence, but only processed cow's milk thereafter. In such a case the Interview should have changed the raw cow's milk consumption to zero at the time of the residence change, but didn't always do so. All of these data entry errors were corrected in the table of milk consumption values.

- The food and milk consumption tables, which contained information about quantities consumed during different time periods for each participant, were searched for overlapping time periods or duplicate change dates. This check was made for each type of consumption (milk, fruit, vegetable, and eggs), as well as for the brands of milk.

- Large changes in consumption of specific food and milk products were examined. A large change was defined to include any increase or decrease by more that a factor of ten in the consumption rate. In addition, to test check for errors that might have arisen from changes in the way a CATI respondent reported consumption rates, e.g., from units per day to units per week, the following additional criteria were defined: (1) a decrease (increase) by more than a factor of 6 in conjunction with a change in reported consumption from units per day to units per week (or from units per week to units per day, respectively), and (2) a decrease (increase) by more than a factor of 4 in conjunction with a change in reported consumption from units per week to units per month (or from units per month to units per week, respectively). Review of CATI audiotapes confirmed that many of these large changes were indeed reported by the CATI respondent. Consumption values that were erroneous due to mistaken entry of the consumption rate were corrected in the table of consumption values.

- Extremely large consumption values were checked for both participant's and mother's diets, for each type of food and milk product category separately, i.e., for glasses, other servings and products of milk for all types of milk consumed, and for raw vegetables, cooked vegetables, raw tree fruit, cooked tree fruit, raw vine fruit, cooked vine fruit, and free range chicken eggs. In addition, the totals for a particular type of consumption, i.e. processed cow's milk, raw cow's milk, total milk consumed, total fruit, and total vegetables were also checked for extremely large values. Descriptive statistics were calculated for each of these separate and combined categories, and records indicating very high consumption levels were identified. The pertinent CATI audiotapes reviewed, and the data were verified or corrected, as appropriate.

- Because consumption of contaminated goat's milk was expected to cause relatively large doses, the audio recordings of all CATIs with an indication of consumption of goat's milk were reviewed and the data verified or corrected as appropriate.

- The CATI allowed the respondent to report that consumption of a given food or milk product changed gradually over a defined time period, for example, from one glass of milk per day at age 6 months to 3 glasses per day at age 3 years. Such "gradual changes" were coded a particular way. All records that were coded to indicate a gradual change, but for which the consumption value did not in fact change, were investigated. Most of these were determined to be correct because the CATI respondent initially

indicated that a gradual change had occurred, but subsequently reported that the quantity consumed had remained constant. In the remaining instances the database was corrected as appropriate, after review of the CATI audiotape is necessary.

- To code a gradual change in consumption of a food or milk product that ended at the end of 1957 (the end of the period for which doses were estimated), the CATI interviewers entered data indicating a change in consumption on 12/31/57 or 12/30/57. (Since dietary data were not collected after that date, this served as a convention to indicate the end of a gradual change.) All other changes of consumption in the year 1957 were investigated to ensure that they weren't meant to indicate gradual changes, and to search for errors in the year of the change date. The CATI audiotapes were reviewed and the table of consumption values corrected as necessary. One CATI Interviewer was found to have consistently recorded gradual changes ending at the end of 1957 incorrectly, using a date other than 12/31/57 or 12/30/57. All the consumption change dates in 1957 in this Interviewer's CATIs that were revised to reflect gradual changes ending at the end of 1957, with the exception of those for which consumption changed either to or from zero, which were considered reliable indicators of nongradual changes.

K.3. Scenario File Construction

To examine the accuracy of scenario file creation, portions of selected scenario files, including the participant's diet and residence history and the mother's diet, were recreated by hand or by a computer program written by someone other than the programmer who created the original scenario files. The participant's diet portion of the scenario file was recreated by hand for approximately 10% of those with CATIs. Two-thirds of those chosen were among the participants with more than the median number of records in the diet portion of the scenario file, and one third from those with fewer than the median number of records. Also included in this group were all participants with a diagnosis of thyroid cancer who had a CATI.

The residence portion of the scenario file was also recreated by hand but for a smaller number of subjects, as this was much less complex than the diet portion. The mapping of residences was also checked for 26 participants (110 places of residence) and no errors were found. In addition, the encoding places of residence within the HEDR domain was tested by checking whether the encoded locations were within the state and county recorded from the CATI or Expanded In-Person Interview. Only 4 errors were found and corrected. There was an error found in the map book, which was prepared by Battelle Pacific Northwest Laboratories for use in determining residence codes. The residence codes in Kittitas county were incorrect as written and subsequently revised. This affected 11 HTDS potential participants, whose residence codes were revised accordingly.

The mother's diet portion of the scenario file was recreated via SAS® programming by a different programmer than the one who created the original diet portion of the scenario file. These two files were then compared.

When the recreations of scenario file data described above identified any discrepancies, the computer programs used to create scenario files were examined for errors of logic or coding, and modified as appropriate.

K.4. Dose Calculation

In order to check the process of scenario file creation and dose estimation, raw dosimetry data for 10 participants whose doses were based on CATI data were provided to an investigator at the CDC. Using that raw data, the CDC investigator created scenario files from those data, and used those scenario files as input to the CIDER program to calculate a set of dose estimates which were then compared to the original estimates calculated by HTDS. Initially, the CDC was not informed of any assumptions made by the HTDS in creating the scenario files, in order to test whether the assumptions she made were similar to those

of HTDS. After a preliminary comparison of the CDC and HTDS dose estimates, it was determined that the CDC used the library of reference diets defined in the HEDR model for persons who consumed commercial ("grocery") milk, while the HTDS used the reference diets for persons who consumed milk from family cows (see section VI.A.3.a below for further description of reference diets). Although the selection of reference diet library had relatively little impact on the estimated doses of participants whose diets were largely if not entirely specified in their CATI data, the grocery milk reference diet library was used by both CDC and HTDS in the subsequent comparisons.

There were several other issues, primarily concerning the handling of unknown or incompletely specified data, for which the CDC and HTDS made different assumptions. These are described in Table V.K-1 below.

Table V.K-1. Differences in Assumptions Used by CDC and HTDS

Issue	CDC	HTDS
Gradual change in consumption of a food, milk, or milk product	The time interval of the gradual change was divided into thirds. Consumption in the 3 resulting subintervals was as follows: First: C(begin), Middle: [C(begin)+C(end)] / 2, and Last: C(end), where C(begin) and C(end) denote the consumption levels at the beginning and end of the interval, respectively.	The time interval was split by year and the consumption in each interval was calculated by successively adding the quantity [C(end) –C(begin)] / (#intervals–1), where C(begin) and C(end) denote the consumption levels at the beginning and end of the interval, respectively.
Unknown food, milk, or milk product consumption quantity	If the quantity was known for some portion of the time, that amount was used to estimate the amount during the time when it was unknown.	Left as unknown (i.e. used CIDER defaults) with the exceptions of 1) other servings of milk or milk products was unknown, or 2) one component of fruit was unknown. For the former, an HTDS default was used, based on tables of median amounts consumed by age, sex, and types of milk consumed. For the latter, the unknown component was set to 0. If more than one component of fruit was unknown, the total was set to unknown.
Combination of milk products and fresh milk	Sum of fresh milk and milk product quantities	Sum of fresh milk quantity plus half of milk product quantity
% Local for vegetables when known for both raw and cooked vegetables	The higher of % local for raw vegetables and % local for cooked vegetables was used	The weighted average of % local for raw vegetables and % local for cooked vegetables was used
% Local for vegetables when unknown for at least one of raw and cooked vegetables	Used 50%	Used 100%
Milk Brands – Brand unrecognized by CIDER	Used brand code	Set to grocer milk
Milk Brands – When brand records start after the milk start date	Assumed first reported milk brands applied from milk start date	Used CIDER default milk brand from milk start date until start of reported milk brands

Despite the differences listed in the table above, the CDC dose estimates were relatively close to those of HTDS, differing by less than 5% for half of the ten, and by less than 20% for eight of the ten; see Table V.K-2 below.

Table V.K-2. Comparison of Dose Estimates by CDC and HTDS

Case	Estimated Dose (Median of 100 Realizations, in mGy) CDC	HTDS	% Difference
1	9.9	10.4	–5.1
2	9.1	9.2	–0.8
3	5.4	10.1	–46.2
4	35.8	25.0	43.1
5	2.5	2.7	–6.6
6	57.7	55.7	3.6
7	94.4	85.0	11.2
8	8.6	8.4	2.9
9	95.9	95.4	0.6
10	12.5	12.2	2.7

The difference in dose estimates for case #3 in Table V.K-2 (5.4 versus 10.1 mGy) was due to a misspecification of the participant's milk consumption in the HTDS estimate; this misspecification was corrected before the participant's final dose estimates were calculated. The difference for case #3 arose from a misspecification of the participant's wean date in the CDC dose estimate. The difference for case #7 resulted from the different methods of combining milk products and fresh milk (see Table V.K-1 above). In summary, this comparison indicated a high level of agreement between dose estimates calculated by HTDS and those calculated by investigators external to the HTDS who were left to devise their own assumptions regarding missing or incompletely specified information.

K.5. Computer Programming

The computer programs that involved major manipulations of the data or complex code other than standard SAS® procedures were reviewed or tested in various ways to ensure they were accurately doing what was intended. The computer programs that created files of outcome data, as well as those that created the files of data regarding the factors analyzed as possible confounding or effect modifying factors were reviewed by a second person. The dose-response programming was checked as follows. First, the Newton-Raphson algorithm used in these programs was written in Pascal. Using small test data files, the output was compared to hand calculations. In addition, a "fixed" data file, with doses and outcomes that would yield a known intercept and dose-response slope, was used to ensure the output was correct. When the HTDS data was used, the fitted values were examined to ensure they were reasonable. The program was first used for simple cases (e.g., one dose realization with one outcome) and then built upon to handle all 100 dose realizations plus three average doses (media, mean, and geometric mean) with multiple outcomes. Once this was completed, the program was written again using SAS® IML, and the output of the two versions compared. Throughout the process, matrix manipulations were performed to ensure they had the properties required of the Newton-Raphson algorithm.

The program to compute an estimated dose from Nevada Test Site exposures for each participant was written by the HTDS Programmer, using Fortran. Output from this program was compared to estimates obtained using the web-based tool provided on the NCI's website (http://rex.nci.nih.gov/INTRFCE_GIFS/radiation_fallout/radiation_131.html) for a small number of participants, to verify that the program and web tool provided the same results.

K.6 Mortality Data

Encoding of causes of death is described in sections V.I.2.d and V.I.3.d above. As a quality control check, text descriptions of causes of death from the death certificates or informant information were compared to the assigned cause of death code for the 543 potential participants for whom cause of death information was obtained. In 13 of 543 (2.4%) cases, the code was revised as a result of this review. In seven (1.3%) of the cases, the code was revised to 410 - Acute Myocardial Infarction from Another, Non-acute Cardiac Condition, in keeping with the coding rules regarding acute cardiovascular disease. In three (0.6%) of the cases, the code was revised from E995 – Injury due to War Operations by Other and Unspecified Forms of Conventional Warfare, to a more specific cause, due to identification of an additional information source on the exact cause of Vietnam war deaths. In one case (0.2%), the cause of death was changed from 770 – Other Respiratory Conditions of Fetus and Newborn to 760 – Fetus or Newborn Affected by Maternal Conditions Which May be Unrelated to Present Pregnancy. The remaining two revisions were due to data entry errors.

The programs written for the mortality analysis were also reviewed. Person-years at risk were calculated by hand and compared to the results of the computer program for selected participants. The numbers of living people in Washington State used in the mortality analysis program for the various sex, age group and calendar year categories were double checked by hand to ensure their accuracy. Similarly, the numbers of deaths in Washington State for specific causes of death and by sex, age group and year of death were also double checked by hand to ensure their accuracy.

VI. Radiation Dose Estimation

A. Background

A.1. Objectives of Dose Estimation

In an epidemiological study concerning a quantitative exposure such as the thyroid dose from Hanford's ^{131}I, the most informative analyses are likely to be those that examine the dose-response relationship in terms of individual measurements or estimates of exposure. When the initial planning of HTDS began, it was anticipated that the Hanford Dose Reconstruction (HEDR) Project, which was then well on its way to completion, would produce a system that could be used to estimate each study participant's thyroid radiation dose. The study design for HTDS was therefore built in part on the assumption that individual dose estimates would be available, although the design was intended to allow the study to succeed in the unlikely event that individual dose estimation was not possible. This assumption had several implications for the study. For example, it implied the need to collect information from which individual dose estimates could be calculated, which led in turn to the CATI component of HTDS. It also implied the need for the HTDS to establish a system that would process such information into a form suitable for use in dose estimation, accomplish the dose calculations, and make the results available for analysis.

The primary objective of this component of the study was to calculate individual estimates of radiation doses to the thyroid for HTDS participants. Specifically the estimates referred to doses to the thyroid from ^{131}I released into the atmosphere from the Hanford site, as calculated by the dosimetry system created by the HEDR Project. Secondary objectives included testing and verifying the accuracy of data that were used for calculation of dose estimates, and the production of data files concerning dose-related characteristics of the study participants, for use by the study statisticians. These data files would include both descriptive data regarding the participants, particularly concerning dose-determining characteristics, as well as data that might be used for alternative characterizations of exposure to Hanford's ^{131}I. An additional secondary objective was added late in the study: to calculate estimates of doses that study participants received from the Nevada Test Site.

A.2. History of the HEDR Project

In 1987 the U.S. Department of Energy directed Battelle's Pacific Northwest Laboratories to conduct the HEDR Project, following the 1986 recommendation of the Hanford Health Effects Review Panel (16). In 1988 a Technical Steering Panel was selected to direct the HEDR project. One of the main objectives of the HEDR Project was to evaluate the feasibility of developing a system to estimate individual radiation doses to the thyroid from Hanford's ^{131}I and if feasibility was demonstrated, to develop such a system. The evaluation of feasibility, often referred to as HEDR Phase I, was completed in 1991. It was concluded that the available data regarding source terms, atmospheric conditions, deposition rates, and environmental and food chain transport were adequate to support the development of a system to estimate radiation doses. It was also concluded that existing models and computer codes could be adapted for estimating doses and analyzing the uncertainty of the estimates. As part of the demonstration of feasibility, a preliminary set of dose estimates was calculated for hypothetical representative individuals in a 10-county area around the Hanford site. These were the dose estimates available at the inception of HTDS and during the development of the HTDS protocol.

Based on the Phase I results, the HEDR project proceeded to develop a dosimetry system that included the capability of estimating radiation doses to the thyroid from Hanford's atmospheric releases of ^{131}I. These dose estimates included contributions from dietary pathways, i.e., the consumption of contaminated milk and food products, from inhalation of contaminated air, and from external exposure. In particular these estimates could be calculated for individuals using specific data regarding residence and

dietary history and other factors. One important part of this phase of the HEDR work was establishing the geographical domain within which doses could be calculated. The resulting domain, roughly 250 miles east to west and 300 miles from north to south, was substantially larger than the 10-county Phase I area. A working version of the dosimetry system was in place by early 1994, and the main HEDR final reports were published in April 1994. Thus an essentially final dosimetry system was available in time for use in calculating doses for the HTDS Pilot Study. This was particularly significant to HTDS, since the final HEDR results differed from the HEDR Phase I results in ways that impacted the design of the HTDS Full Study.

In 1999 and 2000, based on recommendations arising from the National Academy of Sciences review of the HTDS draft Final Report (78) and discussion with HTDS investigators, investigators are Battelle Pacific Northwest Laboratories made a number of modifications in the HEDR model's computer program and data files (see Appendix 22). The resulting version of the HEDR system for dose estimation was used to calculate the dose estimates used for the analyses described in this report.

A.3. Special Challenges of Dose Estimation for HTDS

A.3.a. HEDR Dose Models

The central challenge for HTDS arose from the need to complete a version of the questionnaire and program the CATI before the HEDR model was complete. Thus the data items to be included, and the specific definitions of those items, were not completely known. The most difficult areas in this regard were selection of feeding regimes for family cows, delay times before consumption of certain milk products, milk delivery to homes, identification of dairies inconsistent with HEDR information, definitions of leafy vegetables, and the handling of reference diets.

- Cow feeding regimes were undefined when the CATI was first developed. HEDR investigators initially recommended that the interview ask whether the cows were fed fresh grass or green chop. However in the final dosimetry system, cow feeding regimes were defined by whether or not the cows were grazed on irrigated pastureland. Fortunately the CATI included questions about the source of water for the cows, since it was unknown initially whether water would be a significant source of ^{131}I. The information about water source was used to impute whether the pasture was irrigated.

- The final HEDR model did not allow the specification of milk products other than fresh milk and "stored milk." The difficulty with the stored milk component was that it did not distinguish between relatively fresh milk products such as cottage cheese and ice cream, and products with long intervals to consumption such as aged cheese and canned or powdered milk. HTDS collected information only for milk products that were relatively fresh. To allow for the time lag for consumption of milk products, a conversion factor of 0.5 was applied to these products, after consultation with Battelle investigators. Thus the quantity of fresh milk products was multiplied by 0.5 and then added to the amount of fresh milk to obtain the total amount of fresh milk for use as input data by the CIDER program.

- At the recommendation of HEDR investigators, the CATI included questions asking whether commercially produced milk was purchased at a store or delivered to the home. Since milk purchased in a store might sit a few days on the shelf before being purchased, the difference in "holdup times" could affect dose contributions from the fresh milk pathway by about 20-30%. The design specifications for CIDER included definitions of 22 media containing ^{131}I contamination, including milk categories for grocery milk (purchased) and creamery milk (delivered). However in the final version of CIDER the creamery milk category was not implemented. It was therefore necessary to treat milk that CATI respondents described as delivered to their homes as though it was purchased in stores.

- One important challenge facing HEDR was the reconstruction of the commercial milk distribution system within the HEDR geographical domain. When the HTDS CATI was initially designed, HEDR was able to provide a preliminary list of 55 dairies that operated in the Benton, Franklin and Walla Walla counties. These 55 were included in the HTDS interview materials. The final version of the HEDR system included many more dairies. In October 1995 HTDS received from HEDR investigators a list of 163 dairies that were included in the dosimetry system. Occasionally, of course, CATI respondents identified as a milk source a dairy that was inconsistent with the HEDR data, i.e., that did not serve the area in question at the time in question according to the HEDR data. In 12 instances such inconsistencies were observed in the data from 2 or more CATI respondents. In eight of these 12, the dairy in question was mentioned by only 2 CATI respondents. Information about these inconsistencies was sent to the HEDR Task Completion Working Group and to former HEDR investigators in April 1996. Since the reported discrepancies did not provide definitive evidence of inadequacies in the HEDR commercial milk distribution model, that model was not revised in response to these discrepancies. Therefore HTDS adopted the following approach. Whenever a CATI respondent indicated that the participant consumed milk or milk products from a dairy that did not, according to the HEDR data, serve the area in question during the period in question, the dairy was assumed to be unknown for the participant's dose calculation. This had the effect of assigning the HEDR location- and time-specific default as the source of commercial milk and milk products. If the HEDR model specified that only a single dairy served the location at that time, then that dairy was assumed to be the source of dairy products. If the HEDR model identified two or more dairies that served a region during the time period of interest, then the default was defined as a mixture of milk and milk products from those dairies.

- The definitions of "leafy vegetables" differed somewhat between the HTDS CATI and the HEDR dosimetry system. For example the final HEDR definition of leafy vegetables included string beans, while the HTDS definition did not. Another problem with leafy greens was conversion of servings (the unit used in the CATI) to kilograms (the unit required for input into the CIDER program). The HEDR system did not define how this conversion should be calculated. Therefore, after conferring with HEDR investigators and dieticians, HTDS developed conversion factors based on the weights of servings of individual leafy vegetables.

- Reference diets were built into the HEDR dosimetry system to provide default information about dietary factors. Such default information could be used when all or only part of a participant's dietary history was unknown. In HTDS this occurred whenever the CATI respondent was unable to provide the specific information. The Expanded In-Person Interview given to HTDS participants without CATI respondents included no questions regarding quantities of milk and food products consumed during childhood. Therefore the calculation of dose estimates for those with dosimetry data from the Expanded In-Person was necessarily based entirely on default dietary data. The final HEDR system was limited to a total of four sets of reference diets, each containing 120 combinations of age, sex, lifestyle and season for each of nine categories of food and milk products. The reference diets are defined for four different circumstances: milk from backyard cows, milk from commercial sources, goat milk only and cows fed stored feed only. Nearly every HTDS case fell into one of the first two categories (backyard cow's milk or commercial milk). However the HEDR model did not include reference diets for people who were reported to have consumed unknown quantities of both commercial and family cow's milk. Also the CIDER program allowed for the specification of only a single reference diet in each set of input data. Therefore it was impractical to allow a participant's reference diet category to change over time, and HTDS used the backyard cow's milk reference diet for dose estimation.

A.3.b. Technical Issues

When HTDS began, it was clear that the dosimetry data would be complex, and it was therefore unclear whether a CATI would be feasible. In 1990, the fastest PC had a 386 chip and many on the HTDS

staff were still using 286 IBM-AT computers. The complexity of the data implied that a relational or hierarchical data base structure would be required. Three candidate database management systems were given the most serious consideration. The first, SIR (Scientific Information Retrieval) was originally developed on mainframe computers, but had become available for desktop personal computers. SIR is a hierarchical database that allows more than one record per case all indexed by an ID number. However SIR was relatively inflexible and had very poor data entry features to make it unsuitable for CATI. Two relational database management systems, ORACLE and INGRES, originally developed for mainframe environments, were also available on personal computers. Of these two, INGRES was selected on the basis of its flexibility, superior data entry capabilities, substantially lower cost, availability of a local office with technical support, and ability to run on relatively modest personal computers. The flexibility, stability and features of INGRES allowed it to meet all of the study's needs. A copy of the SIR product was also required by the study for use in processing dosimetry data through several steps required to create input files for the CIDER program.

Flexibility of the dosimetry data base management system was important since the CATI was modified several times after data collection began. Moreover the INGRES-based system allowed the capability for interviewers to revise responses in real time during the CATI. Designing the system to permit interviewers to return to and modify previous responses during the interview presented many challenges. However it was considered important since it would minimize the need to temporarily discontinue interviews to permit entry of revisions that would impact the appropriateness of subsequent questions. Since the CATI was expected to be a significant imposition on the time and altruism of the respondent, every effort was made to minimize the number of temporary discontinuations.

The ability to correct data after the interview was also an essential component of the dosimetry data base management system. A separate set of data entry programs were written exclusively for data correction. Initially, the Systems Analyst was the only person allowed to make data corrections. After about one year much of this responsibility was shifted to the CATI Interviewers, who by then had enough experience with the data to make many kinds of corrections, and to judge when a correction was so complex or unclear that it had to be performed by the Systems Analyst. In such instances the Interviewer completed a data correction form and the Systems Analyst made the corrections.

A.3.c. Logistics

HTDS had three computers available for CATIs, each with its own copy of the CATI database. Each CATI database contained only data from the interviews conducted on that computer. After the CATI was completed and any necessary corrections made to the data, the Interviewer copied the data to the local area network maintained by the Epidemiology Program of the Fred Hutchinson Cancer Research Center. Data from the three CATI databases were then captured from the network and combined into a single database on the Systems Analyst's computer.

Creating scenario files (i.e., input data files required by the HEDR dosimetry system) from the CATI databases was a complex process performed by the Systems Analyst. It involved merging records for 18 components of the participant's diet that could vary independently over time into a single set of sequential records. Each diet component existed as a separate table in the INGRES CATI database. They were combined into a single table in the SIR database using INGRES's report writer, programs written in FORTRAN, and the SIR programming language. These were then merged with data regarding the participant's residence history, birth date, and mother's diet if necessary to create the scenario file using a FORTRAN program. For participants whose doses were based on data from the Expanded In-Person Interview, the procedure for creating scenario files was similar but somewhat simpler, because that interview did not collect information about quantities of food and milk products consumed by the participant.

B. Dose Estimation Procedures

When the HTDS protocol was developed in 1993, plans regarding the methods for calculating doses to the thyroid from Hanford's atmospheric releases of ^{131}I could not be specified, since relatively little was known about the dosimetry system that would be available. It was assumed that a dosimetry system would be available, and that it would be capable of calculating doses for HTDS participants using the data collected in the CATI or Expanded In-Person Interview. It was highly likely, though not certain, that dose calculations would be performed by some agency other than HTDS.

B.1. Staffing and Logistics

The study's Systems Analyst/Programmer had primary responsibility for the calculation and management of dose estimates. This included the following tasks: developing procedures for capturing and processing CATI and Expanded Interview data into a format suitable for dose calculations (scenario files), transferring data to the custodian of the dosimetry system to have the calculations performed, receiving the dose estimates back from the custodian of the dosimetry, and making the dose estimate data available to HTDS investigators for statistical analysis.

The HEDR model originally used by HTDS for was installed on a Sun workstation administered by the CDC in Atlanta. Scenario files were created in Seattle and transmitted to Atlanta via the Internet. This version of the model was used to calculate doses used in the analyses for draft HTDS Final Report.

Scenario files typically contained data for between 40 and 45 participants, and the Sun installation of CIDER typically required about 70 minutes to calculate doses for those participants. After the doses were calculated by CIDER, HTDS transmitted the results to Seattle, again via the Internet. CIDER computed 100 realizations of dose for each year from 1944 to 1957 (14 years), and for each of the 10 pathways (inhalation, external exposure, and ingestion pathways for eight food and milk categories). This resulted in at least 14,000 realizations of dose for each participant. If the participant moved during a year, 100 realizations for each of the 10 pathways were computed for each location during the year. HTDS wrote a program to combine the 14,000 realizations into 100 realizations of total dose.

B.2. Revisions of the HEDR Model and Computer Programs

The HEDR model, and more specifically the CIDER program, were revised a number of times during the course of HTDS. A number of HTDS suggestions were incorporated in the final version of CIDER:

1. The maximum number of sources of fresh milk was increased from 3 to 5. Many households had multiple sources of milk such as a backyard cow, commercial milk delivered to the home or purchased at a store, and milk served at school.

2. Goat's milk was retained in the final HEDR model. HEDR investigators considered dropping goat's milk, however HTDS CATI data showed that about 1 % of households drank goat's milk.

3. Cow feeding regime #4 (cow fed mostly stored feed) was retained in the final HEDR model. CATI data showed there were cows fed entirely with hay and stored feed.

4. The maximum number of diet specifications was increased (to 860). This was necessary to accommodate the multiple changes in diet that were typical of HTDS participants. HEDR initially set this limit much lower, based on an assumption that many individuals would share common diets. Even with this increased limit, however, only 40-45 CATI cases could be processed at one time.

5. In the summer of 1997 HTDS detected an error in how the CIDER program handled breast-feeding of participants and helped identify and test the correction

In response to suggestions made in the National Research Council's (NRC's) review of the draft Final Report of the HTDS (78), a number of further revisions were made in the CIDER program by investigators at Battelle Pacific Northwest Laboratories. These are described in detail in Appendix 22, which reproduces a letter report produced by investigators at Battelle Pacific Northwest Laboratories (126). One of the most important revisions was to provide HTDS a version of the CIDER program and the related data libraries that could be run on a desktop personal computer with relatively modest memory (see section 7.1 of Appendix 22). This eliminated the need for a cumbersome procedure described above for passing scenario files and dose estimation output between the HTDS offices in Seattle and the CDC in Atlanta. The HEDR model used by HTDS to estimate the doses used in this report was run on an IBM compatible PC located in the HTDS office. The doses were calculated by the HTDS programmer in much the same manner as described above, although the time CIDER typically required to calculate doses for 40-45 participants decreased to about 7 minutes.

The Battelle investigators made two other significant revisions in the CIDER program. As described in section 7.2 of Appendix 22, the handling of uncertainty in dose conversion factors (DCFs) was revised. The HEDR model accounted for these uncertainties by generating 100 realizations of each age- and sex-specific DCF according to defined uncertainty distributions (127). In the original implementation of CIDER, the order of these realizations was fixed. That is, for every participant, the first dose realization was calculated using the first realizations of the DCFs, the second dose using the second DCF realizations, and so on. This created an artificial correlation between dose estimates of different participants, since they all shared common values of the DCFs in each realization. The revised version of CIDER therefore included an option to randomly permute the order in which the 100 DCFs are selected for each participant, thereby eliminating the artificial correlation. This option was employed for all dose estimates used in the analyses described in this report.

The second significant revision in the CIDER program provided options to assign uncertainties to participant-specific dietary consumption data obtained from CATIs (see section 7.3 of Appendix 22). In the original implementation of CIDER, quantities of food and milk products consumed by a person were treated as uncertain only if they were specified as unknown in the scenario file of input data. If the amount of a food or milk product that a person consumed could be specified in a scenario file, then CIDER treated that amount as fixed, with no uncertainty, in estimating the resulting dose. For most HTDS participants with doses based on CATI data, age-specific quantities of foods, milk, and milk products consumed were reported by the CATI respondent. While it was recognized from the beginning of HTDS that it is unrealistic to ignore the uncertainties in dietary data collected from interviews several decades after the exposure period of interest, the original version of the CIDER program provided no practical means to incorporate that uncertainty. The final version of CIDER includes options to assign uncertainties to reported dietary intakes. These options are described in detail in section 7.3 of Appendix 22

The final version of CIDER and the related data libraries included two other revisions that had only limited impact on the dose estimates. These included correction of source terms beginning August 1951 (see section 3.1 of Appendix 22), and of the uncertainty distribution of fetal dose conversion factors (section 6 of Appendix 22).

C. Doses from the Nevada Test Site

Information released by the U.S. National Cancer Institute (NCI) shortly before and during October, 1997, indicated that persons living in the contiguous 48 states during the 1950s and 1960s were exposed to various levels of ^{131}I released from the Nevada Test Site (NTS). The material released by NCI included estimates of dose for various representative individuals for all counties in the 48 states, as well as more detailed data regarding estimated dose by shot (i.e., by individual test detonation), county, and age. Limited preliminary comparisons for HTDS participants suggested that in many cases the reported NTS

dose estimates were comparable to or even greater than the estimated Hanford doses. Therefore it was judged necessary to add exposure to ^{131}I from the NTS to the list of potential confounding factors.

The CATI and In-Person Interviews included complete residence histories for all participants for the period from December 1944 through 1957. For periods when a participant lived outside the HEDR geographical domain, the county and state of residence were recorded, although details regarding diet and sources of food and milk were not obtained. This was fortuitous, since it provided a means to calculate estimates of NTS-derived dose. Using data regarding representative doses by age and county available from the NCI's website, the HTDS Systems Analyst/Programmer created a program that calculated estimated NTS doses for study participants, based on their residence histories through 1957.

VII. SPECIAL CONSIDERATIONS

A. Assessment of the Feasibility of a Health Study in Native American Populations

A.1. Background

Nine Native American tribes and nations have reservations and ceded lands in the region around Hanford: Colville, Couer d'Alene, Kalispell, Kootenai, Nez Perce, Spokane, Umatilla, Warm Springs, and Yakama. Members of these tribes and nations were exposed to ^{131}I from Hanford, and the original Congressional mandate that led to the HTDS called specifically for the inclusion of "Indian tribes and tribal organizations."

The approach taken in the HTDS regarding the Native American communities was determined by two important characteristics of those populations. First, the lifestyles of many Native Americans were quite different in many respects from those of the non-Native population. In particular, many Native Americans followed traditional cultural practices, especially regarding diet and sources of foods, which might influence the doses they received from Hanford's ^{131}I but which were not explicitly modeled in the HEDR calculational programs. Moreover, many Native Americans maintained a seasonal migratory pattern of residence. Second, because the tribes and nations have sovereign rights recognized by the United States, conduct of a research project such as HTDS would require the approval of each tribal government and active cooperation of tribal members to obtain culturally sensitive data.

As stated in the HTDS protocol (1), the objective of the HTDS with respect to the Native American populations was to assess the feasibility of conducting a study to determine whether thyroid disease was increased among Native Americans exposed to ^{131}I from Hanford. The approach taken to meet this objective involved the following steps:

1. Identifying study designs that could meet the main objective, i.e., to determine whether thyroid disease has increased among Native Americans exposed to ^{131}I from Hanford.

2. Establishing guidelines for assessing whether any of the proposed designs had adequate probability of providing a definitive conclusion regarding the main objective. These guidelines were to be established in collaboration with CDC staff and representatives of the tribes involved.

3. Analyzing demographic data and estimates of thyroid doses from ^{131}I, using dietary and lifestyle information provided by the tribes, in relation to the established guidelines to reach a conclusion about feasibility of a study in the Native American population.

These activities were undertaken in parallel with those of the Full Study. The sections below briefly describe the progression of this component of the HTDS. Demographic data and information about lifestyle practices collected by each tribe, as well as radiation dose estimates specific to each tribe, are considered proprietary and belong to the tribes. These data were made available to the HTDS investigators for purposes of assessing the feasibility of a study in the Native American population with the understanding that they would not be disclosed. Therefore, no data specific to individual tribes are included in this report.

A.2. Initially Recommended Study Design and Guidelines for Assessing Feasibility

Since the main objective of a study in the Native American population is the same as that for the HTDS Full Study, the choice of possible study designs was subject to the same constraints as the HTDS Full Study (see section IV.A.1 above for a discussion of study design considerations). Therefore, the

HTDS investigators initially recommended that a retrospective cohort design using individual dose estimates, similar to that used for the HTDS Full Study, would be most appropriate for a study in the Native American population.

To begin the feasibility assessment of conducting such a study, it was necessary first to obtain information from each tribe about the number of persons who might be available and willing to participate in a study It was necessary to obtain information to estimate thyroid radiation doses that members of the tribe would have likely received from Hanford. When the HTDS was initiated, work was already underway in conjunction with the Hanford Environmental Dose Reconstruction Project to begin to collect such information within the tribes. A working group was formed to facilitate this effort, and to provide technical assistance. This working group was composed of representatives of each tribe, the Technical Steering Panel of the HEDR project, Battelle, Pacific Northwest Laboratory, the CDC, and the health departments of the states of Washington, Oregon, and Idaho. The HTDS joined this group (the Native American Working or NAWG), and was represented by one of the study investigators at each meeting. This provided a close link between the HTDS and each tribe throughout the entire process of data collection and dose estimation for the Native American population.

As the data collection effort proceeded, it became clear that guidelines for assessing the feasibility of an epidemiologic study should be developed and agreed upon *prior to* examination of any tribal-specific data. At a meeting of the Native American Working Group in January 1994, HTDS investigators proposed the following guidelines for assessing whether an epidemiological study of the recommended (retrospective cohort) design should be conducted in the Native American population.

- Justification of such a study would require pilot data indicating the feasibility of identifying and recruiting adequate numbers of people with a range of radiation doses sufficient to ensure that a one-sided test at the 5% critical level has at least 80% power to detect a linear dose-response for the probability of thyroid neoplasia with a slope of 10^{-5} per mGy.

- If pilot data indicated that such a study would have substantially less power, e.g., below 70%, to detect an effect of this magnitude, then a study in the Native American population would not be recommended on scientific grounds.

This criterion was analogous to that initially proposed for the decision about whether to proceed with the HTDS Full Study. In particular, the target of 80% power to detect an effect of 10^{-5} per mGy was considered scientifically sound, providing a sufficiently high level of statistical power (80%) to detect a relatively small effect (10^{-5} per mGy). It was also considered achievable, based on the dose data available at the time, i.e., the HEDR Phase I dose results.

As a result of the HTDS presentation at the January 1994 meeting, the NAWG formed a Subcommittee on HTDS Study Design to further evaluate the proposed guidelines. In May 1994, this Subcommittee of the NAWG requested the HTDS investigators provide a document regarding possible study designs that might be considered for a thyroid study in the Native American population. In June 1994 the requested document was submitted to the Subcommittee (128). In that document, the HTDS investigators concluded that, given the objective of determining conclusively whether thyroid disease was increased among Native Americans exposed to Hanford's ^{131}I, the most appropriate study design remained a retrospective cohort study with individual dose estimates, similar to that of the HTDS Full Study.

A.3. Modified Guidelines for Assessing Feasibility

When the essentially final HEDR results became available in April 1994 (14, 118), it was apparent that the range of dose estimates for HTDS participants would be substantially smaller than the HEDR Phase I results had suggested. Consequently, the preferred criterion of having 80% power to detect an effect of 10^{-5} per mGy no longer appeared achievable. Therefore, modification of the recommended

criterion had to be considered to ensure adequate power (at least 80%) to detect a dose-response effect of 5×10^{-5}/mGy for thyroid neoplasia. Although this represents a substantial decrease in the power of the study, it was still considered scientifically justifiable (see section V.A.5 above). For example, based on projected baseline probabilities for thyroid neoplasia of 5% and 2% for women and men, respectively, in the HTDS cohort, an effect of 5×10^{-5}/mGy corresponds to doubling the probability for women at a dose of 1000 mGy, and to approximately tripling the risk for men at that dose. This is roughly the magnitude of effect seen in the study of persons in Utah exposed to fallout from the Nevada Test Site (60).

At an October 1994 meeting of the NAWG, representatives of the tribes and nations agreed to provide demographic information and representative dose estimates that were calculated for them by staff at Battelle Pacific Northwest Laboratory as part of the Phase I Native American component of the HEDR project. From 1994 through early 1996, six tribes provided the estimated dose data to HTDS. Five of these tribes also provided information about their numbers of members during the 1940s and 1950s. Based on these data, HTDS investigators calculated predictions of the dose distributions, cohort sizes, and statistical power that might be available for a study in the Native American population.

In May 1996, at a meeting of the Intertribal Council on Hanford Health Projects or ICHHP (which had by then taken the place of the NAWG), HTDS investigators made a presentation regarding the impact of the final HEDR results. For the reasons discussed above, they recommended that the guidelines for assessing feasibility be relaxed. They also presented the results of the power calculations based on the preliminary dose data that had been provided by six tribes or nations. These calculations showed that the projected range of doses and cohort sizes were not large enough to meet the modified guideline of 80% power to detect an effect of 5×10^{-5}/mGy for thyroid neoplasia. However it was also recognized that the dose data available at that time were quite limited with respect to both the number of interviews conducted with tribal members, and the number of tribes completing data collection. Thus, these data would not likely provide sufficiently accurate projections of the dose distributions that would actually be obtained if a Native American study were performed. Consequently, the HTDS investigators recommended that the final determination regarding feasibility of a study be postponed until the second stage of data collection and dose calculation for the tribes was complete.

A.4. Final Assessment of Feasibility

The second stage of data collection took place during 1996 and 1997. One tribe did not complete this stage of data collection and dose estimation. This data collection was intended to provide input data for the calculation of estimated doses for hypothetical representative persons based on realistic assumptions about diet, food sources, and seasonal changes in residence. In 1997 and 1998 CDC staff communicated with each tribe or nation to obtain approval of the assumptions used for calculation of these representative dose estimates. Between late 1997 and mid 1998, CDC and HTDS staff calculated representative dose estimates as data and approvals became available from the tribes.

Dose estimates for each tribe were calculated using several scenarios for the hypothetical representative persons. These scenarios were defined by the following factors:

- Sex.
- Year of birth, including at least 1940, 1942, 1944, 1945, 1946 for every tribe. For three of the tribes, earlier and later birth years were also included. The date of birth was assumed to be January 1 in each year.
- Age at weaning, including at least 0 months (indicating absence of breast-feeding) and 12 months. Older ages at weaning were included for the calculations of some tribes.
- Diet, traditional diet as reported by the tribe versus reference diet as provided by the CIDER program.

Examination of the representative Native American doses revealed that the pattern of estimated dose in relation to birth year was generally the same as that of the individual dose estimates of the HTDS Full Study. That is, estimated doses tended to be highest for representative persons born in 1945. The representative doses decreased with decreasing (earlier) birth year, and were also lower for the 1946 births. One tribe's dose estimates differed slightly from this pattern in that the highest representative dose estimate was for a 1944 birth. However doses for 1945 were similar to those for 1944, and doses still decreased as one moved away from 1944 or 1945, toward either earlier or later birth years.

These representative dose estimates were used to perform statistical power calculations in essentially the same way as those presented by HTDS investigators to the May, 1996, meeting of the ICHHP. Note that the new power calculations differed from the earlier calculations in the following respects: 1) they were based on the representative dose estimates calculated in 1997 and 1998 and based on presumably more accurate scenarios for tribe-specific dietary data and residence histories, and 2) they were based on eight of the nine tribes, rather than the six tribes in the earlier calculations (five tribes provided data for both sets of power calculations).

The statistical power of a study depends in part on the mean and variance of the distribution of doses that would be estimated for the study participants (see Appendix H in HTDS Protocol [1]). These quantities were estimated by 1) estimating the mean and variance of the dose distributions and number of participants for birth year cohorts within each tribe, and then 2) calculating the mean and variance of the overall dose distribution that would result.

To assess the feasibility of a study in the Native American population, initial calculations of statistical power were performed using nonconservative assumptions, i.e. assumptions that would tend to produce an overestimate of the statistical power. This was done as a scoping calculation: if the overestimated statistical power was too low to justify conduct of a study, then the even lower projections of power that would result from using more realistic assumptions would also be evidence against feasibility. If, on the other hand, the initial scoping calculations indicated that adequate statistical power might be obtained, then more careful evaluation of the projected statistical power would be pursued. The non-conservative assumptions that were made for the initial scoping calculations were as follows.

Assumption 1. The projected mean of the dose distribution was calculated by assuming that, within each tribe, the mean dose that would be obtained for each of the 1940, 1942, 1944, 1945, and 1946 birth cohorts would equal the maximum representative dose calculated for that tribe and birth year.

Assumption 2. For each tribe's 1941 and 1943 birth cohorts, for which representative doses were not calculated, the mean was assumed to equal the maximum representative dose calculated for that tribe's 1942 and 1944 birth years, respectively.

Assumption 3. All nine tribes and nations would participate in a proposed study of the Native American population, even though one tribe did not participate in the second stage of data collection and representative dose calculation. Since representative dose estimates were not available for this tribe, it was assumed that each of its birth cohorts would have the same mean dose as the tribe with the highest representative dose estimates. This assumption is quite non-conservative, since the mean doses for this tribe would almost certainly be much smaller.

Assumption 4. For each birth year cohort within each tribe, the variance (V) of the doses was assumed to equal the square of the mean dose (M),

$$V = M^2 .$$

The representative dose calculations provided estimated doses for certain types of individuals, but did not provide estimates of the variance that might be observed in a population of real individuals. Therefore, the relationship between mean and variance of doses for populations of real individuals was estimated from the individual dose data available from the HTDS Full Study. The 3191 living evaluable

in-area participants in the Full Study were divided into 100 subgroups defined by sex, year of birth, and geostratum, and the mean and variance of the dose estimates for each subgroup were calculated. Regression analyses indicated that the relationship between variance and mean was approximately of the form

$$V = M^B .$$

For dose estimates which used individual residence histories, individual information collected by the CATI, and HEDR default values for items for which CATI data were not available or for individuals without a CATI, the exponent B was estimated to be 1.8 ± 0.03 (S.E.). For dose estimates which used individual residence histories and HEDR default values exclusively (i.e., no individual CATI data), the exponent B was estimated to be 1.7 ± 0.03. Using an exponent of 2 results in larger estimates of variance, and therefore higher projections of statistical power, than would be obtained with 1.7 or 1.8.

Assumption 5 Non-conservative assumptions were made regarding the numbers of participants who would be available from the nine Native American populations. In particular, for tribes that provided detailed demographic data, it was assumed that all members of the included birth cohorts would be living evaluable participants. For all other tribes it was assumed that a total of 1000 living evaluable participants would be available. This constitutes perhaps quite an overestimate of the number of participants, given the relatively small size of several of the tribes.

Based on these assumptions, sample sizes and dose distributions for a Native American study based on the 1940 – 1946 birth year cohorts for all nine tribes were projected, and the resulting statistical power was calculated. For the initial scoping calculations, a sample size of 6426 living evaluable participants was projected. For thyroid neoplasia, assuming the same background rates as were assumed for the planning of the Full Study, i.e., 5% for women and 2% for men, there would be only 50% power to detect a dose-response effect of 5×10^{-5}/mGy.

In addition to sample size and the mean and variance of the dose distribution, power is also influenced by the baseline probabilities of disease. In particular, all other factors being equal, power increases as the baseline probabilities decrease. Therefore, the sensitivity of the estimated power to the assumed background rates was investigated as part of the initial scoping calculations. To provide a rather extreme boundary for estimated power, calculations were repeated assuming that the baseline probabilities of thyroid neoplasia in the Native American population are only *half* of those assumed for the Full Study (2.5% rather than 5% for women, 1% rather than 2% for men). Under this assumption there would still be power of only 71% to detect an effect of 5×10^{-5}/mGy. Unfortunately, there are no good estimates available of the baseline prevalence of thyroid neoplasia among the nine tribes in the Hanford region. Thus, the assumption of one half used above is intended only to provide a wide boundary of what might be achieved in study power. It is not based on specific estimates of disease prevalence in the Native American population.

In summary, initial sample size and power calculations were carried out based on data provided by eight of the nine tribes under consideration. It is presumed that these data reflect lifestyle patterns and practices specific to each tribe, and that therefore the representative dose estimates more accurately approximate the dose members of each tribe would have likely received from Hanford than earlier estimates. Similarly, it is presumed the demographic data provide a reasonably accurate estimate of the size and demographic makeup of each tribe around the time of the Hanford releases. The five assumptions described above that form the basis for the scoping calculations are deliberately non-conservative. Within a reasonable framework, they are intended to err in the direction of overestimating possible doses, variance of doses, numbers of available participants, and members of participating tribes. Even under such extreme assumptions, a study nearly double in size as the HTDS Full Study (6426 living evaluable participants) would have only 50% power to detect an effect of the magnitude considered scientifically sound. Even under the more extreme assumption that the baseline probabilities for thyroid neoplasia are only half of

those assumed in the Full Study, a study of 6426 living evaluable participants would only have 71% power to detect the same magnitude of effect.

Based on these results, the HTDS investigators recommended that it was not feasible, nor scientifically justified, to undertake a study of the same design as the Full Study (i.e., a retrospective cohort study). Such a study would require more than 6400 living evaluable Native American participants, and would have at most 50% power to detect a dose-response effect of $5x10^{-5}$/mGy for thyroid neoplasia.

B. Coordination with the Advisory Committee

In June of 1990, an Advisory Committee was appointed by the Secretary of the Department of Health and Human Services to advise and consult with the CDC regarding the design and conduct of the study. The committee was established pursuant to the *Federal Advisory Committee Act, 5 U.S.C. (Appendix 2)*. The role of the committee was to review the development of the study protocol and conduct of the Pilot Study, assist in determining the feasibility and design of a full-scale epidemiologic study, and advise the CDC on the analysis of the study data.

The committee was to be made up of seven scientific and lay members representing different areas of expertise or knowledge. The original members appointed to the committee were: 1) Mr. Lou Stone, representing Native Americans; 2) Dr. Owen Hoffman, Ph.D., representing expertise in Radiation Science; 3) Dr. Genevieve Matanowski, M.D., Dr. P.H., representing expertise in Epidemiology; 4) Mr. Jim Thomas of the Hanford Education Action League, representing environmental organizations in the Pacific Northwest; 5) Dr. Arthur Schneider, M.D., representing expertise in thyroid disease; 6) Ms. Christine Holmes, representing the people of Washington State; and 7) Dr. Larry Jecha, M.D., Health Officer for Benton-Franklin Counties, ad hoc member. The first meeting was held in Atlanta, Georgia in March 1991. Dr. Jecha was appointed Chairman by the CDC.

Prior to the first meeting, Ms. Holmes notified the CDC that she would be unable to participate as a member of the committee. Her position was replaced with Ms. Kristine Gebbie, Secretary of Health for the State of Washington. Due to concern that the affected population might not be adequately represented by a state official, the committee requested a consultant position be added to the committee. This would be a non-voting member familiar with the concerns of those who felt their health had been affected by radiation from Hanford. Ms. Pamela Hoefer, R.N., of the Hanford Downwinders Coalition was selected to fill this position. Ms. Hoefer attended meetings from March 1992 to February 1993.

Over the course of the study, several individuals were replaced as their term of service expired, Dr. Maureen Hatch (epidemiology) replaced Dr. Matanowski in August 1995, Dr. Marlene McKetty (dosimetry) replaced Dr. Hoffman in August 1995, and Ms. Elizabeth Ward (State of Washington) replaced Secretary Gebbie in August 1995. Ms. Judith Jurji replaced Ms. Hoefer in October 1993 as a consultant to the Committee representing the Hanford Downwinders Coalition.

Initially, meetings of the committee were to be held on a quarterly basis in Atlanta. In recognition of the interest in the Pacific Northwest in such proceedings, however, the committee asked that at least one meeting per year be held in Washington State. Following completion of the Pilot Study, meeting frequency was reduced to approximately once per year, with the majority of these held in Seattle, Washington.

Meetings of the Advisory Committee were uniformly open to the public. All materials presented to the committee became public record, with copies available for members of the public at the meetings. Time for public comment and questions was allowed in each meeting's agenda. In addition, meetings held in Washington State were usually accompanied by an evening Public Meeting to allow members of the public to attend and ask questions regarding the study.

Each meeting of the Advisory Committee began with an update on the progress of the study since the previous meeting. These presentations included the status of preparations for the field work, or later, the numbers of study participants completing each phase of the study. Updates on the separate work concerning Native American populations were also included. In addition to monitoring ongoing operations, the Committee focused most of its attention on the following items: 1) review and approval of the initial study protocol; 2) review of the Pilot Study Final Report and recommendations to move forward with a Full Study; 3) development of guidelines for assessing the feasibility of a study of Native American populations; 4) review of the Analysis Plan for the Full Study; and 5) review of the Communications Plan for the Full Study.

C. Public Information

An important aspect of this research was the provision of prompt, accurate, and complete information to the public. In this context it was crucial that contacts be established with members of the populations most interested in (and potentially affected by) the work. Interested parties included representatives of the States of Washington, Oregon, and Idaho, the Native American Tribes and Nations in the study areas, and local area residents.

The public information activities of the study were designed to accomplish the following goals:

1. To assure that residents of the region understood the issues that led to the initiation of the study, the purpose and objectives of the study, its basic epidemiologic design, and the time schedule within which it was to be conducted.

2. To provide opportunities for the public to express concerns and comments regarding the design and conduct of the study, and to answer public questions regarding all aspects of the project.

3. To create public interest and support for the study, particularly in ways that might enhance participation by persons selected to be study participants.

4. To assure broad dissemination and proper interpretation of final study results.

Although all members of the Study Management Team fully expected to contribute in an effort to keep the public informed and to answer questions, Dr. Scott Davis assumed primary responsibility for coordinating such activities. An important initial step in the overall approach was to establish contact with counterparts on the TSP responsible for public information activities (the Communications Subcommittee, chaired by Ms. Mary Lou Blazek). Thus, while the HEDR Project was still underway, the two projects coordinated their efforts to keep the public as well as agencies of the states and Native American Tribes and Nations well informed regarding the planning and the progress of the study. This process was greatly facilitated by the fact that one of the HTDS investigators, Dr. Kenneth Kopecky, was also a member of the TSP and served on the Communications Subcommittee.

Throughout the HTDS, and particularly in its early phases, the SMT participated in public meetings held during the bimonthly meetings of the TSP, and contributed to the planning activities of the Communications subcommittee of the TSP. Members of the SMT, or the Project Manager, attended each TSP meeting. In addition, a member of the HTDS staff attended all meetings of the Communications Subcommittee. In an effort to work more extensively with the TSP in the area of providing public information, at least one member of the SMT was present whenever possible at all TSP-sponsored public meetings and workshops. The HTDS also supplied the TSP with a Fact Sheet that was included with TSP fact sheet mailings. This written material was updated periodically as the study progressed.

Several separate approaches were also taken to provide information to the public regarding the HTDS. Initially, the study protocol was made available for public review and comment prior to its

submission to the CDC and the Advisory Committee. In conjunction with this activity, a series of public (town) meetings were held throughout the Northwest to discuss the protocol with the public and to answer specific questions. Similar public meetings were held in conjunction with meetings of the Advisory Committee held in the Pacific Northwest.

In addition to the study Fact Sheet mentioned above, several study brochures were developed and a newsletter describing the progress and status of the study was initiated. A master mailing list, which included the lists previously maintained by the FHCRC, the CDC, and the HEDR Project was assembled to mail the newsletter and brochures to interested individuals.

Finally, study investigators and staff were available to answer questions on a regular basis. A phone line was designated in the Seattle study office for public inquiries, and a toll-free telephone number was established at the Fred Hutchinson Cancer Research Center for the Hanford Thyroid Disease Study (1-800-638-HTDS). Persons selected as study participants were encouraged to use the toll-free number to contact the study office if they had questions or scheduling conflicts. The toll-free number was also made available to the general public so that anyone with questions or comments could easily contact the study. As access to the World Wide Web via the internet became more common, a web site for the study was established at the FHCRC. All study brochures and newsletters have been available at that site since January 1997, and are updated as appropriate. Links to the FHCRC, Centers for Disease Control and Prevention, and Hanford Health Information Network sites have been established. The HTDS web site can be accessed at http://www.fhcrc.org/science/phs/htds.

VIII. STATISTICAL METHODS

A. General Approach

A.1. Objectives of the Statistical Analysis

The primary objective of the HTDS was to determine whether thyroid disease has been increased among persons exposed to radioactive iodine released from the Hanford Nuclear Site between 1944 and 1957 (see section III above). To meet this overall objective, the statistical analysis had the following three specific objectives:

1. To estimate, and test the statistical significance of, exposure-response relationships between various thyroid disease outcomes (and other outcome and response variables) and measures of exposure (dose) to radioactive iodine from Hanford.
2. To identify and analyze the effects on these dose-response relationships of any confounding or effect-modifying factors.
3. To investigate, to the extent possible, the shapes of any dose-response relationships that are found.

These specific objectives are discussed in more detail in the following three sections.

A.1.a. Estimation and Testing of Dose-Response Relationships

The primary analyses of this study examined dose-response relationships for the following response variables:

1. Thyroid disease outcomes
   - Thyroid cancer
   - Benign thyroid nodule
   - Thyroid neoplasia
   - Any thyroid nodule (benign, malignant, or suspicious for follicular neoplasm)
   - Hypothyroidism
   - Autoimmune thyroiditis (Hashimoto's thyroiditis)
   - Graves disease
   - Autoimmune thyroid disease (i.e., Hashimoto's and/or Graves)
   - Hyperthyroidism
   - Multinodular thyroid gland
   - Simple goiter
   - Other thyroid disease

2. Other outcome variables
   - Hyperparathyroidism
   - Ultrasound-detected abnormalities of the thyroid (thyroid UDAs)

3. Other response variables
   - Thyroid stimulating hormone (TSH)
   - Total thyroxine (T4)
   - Triiodothyronine resin uptake (T3RU)
   - Free thyroxine index (FTI)
   - Anti-thyroid anti-microsomal antibody (AMA) or anti-thyroid peroxidase antibody (anti-TPO)
   - Anti-thyroglobulin antibody (anti-TG)
   - Thyroid mass
   - Serum calcium

The list of thyroid disease outcomes was comprehensive since the objective ("to determine whether thyroid disease is increased ...") included all thyroid diseases. Therefore it included thyroid diseases for which associations with ionizing radiation have been reported in other settings (thyroid cancer, any thyroid nodule, autoimmune thyroiditis, and hypothyroidism (see section II.B above), as well as other diseases for which associations have not been reported. In view of the public concern about possible unanticipated effects of exposure to ^{131}I from Hanford, and of the Congressional mandate that the study address thyroid morbidity, the exposure-outcome relationship was analyzed and reported separately for each of the outcomes listed above. While the various outcomes can be distinguished in terms of the quantity and strength of the existing evidence for association with exposure to ^{131}I, such distinction played no role in determining how or how extensively the various outcomes were analyzed. Similarly, while the outcomes might be distinguished in terms of severity of impact on a person's life, the same level of effort was expended to assess each diagnostic outcome and its relationship to ^{131}I exposure.

In the primary dose-response analyses, the exposure for each individual was represented by the estimated radiation dose to the thyroid from ^{131}I, as calculated using the CIDER program created by the Hanford Environmental Dose Reconstruction Project. CIDER calculates a dose estimate only if the participant resided within the 246-by-306 mile HEDR geographical domain after December 25, 1944(129, 130). Therefore ad hoc estimates of the thyroid dose were used for study participants for whom CIDER did not produce a dose estimate; see section VIII.C.1.a.3 below for further details.

The primary dose-response analyses for disease outcomes were based on regression models in which the probability of having the outcome of interest varies as a linear function of thyroid dose. In particular, this primary model permitted background probability of the outcome (i.e., the intercept parameter) to depend on sex, but assumed a common regression coefficient (slope) for dose. The regression coefficient can be interpreted as the change in the probability of the disease outcome, per unit change in dose. So, for example, a slope of 0.005 per Gy indicates that the probability increases by 0.005 (i.e., five per thousand or 0.5 percentage points) for each dose increase of 1 Gy and a slope of 0 per Gy indicates that the probability does not change with dose. Estimation of the dose-response relationship was accomplished by estimating the slope of this stratified linear dose-response model. Since the purpose of the study was to determine whether thyroid disease has been increased, significance testing focused on the null hypothesis that the probability of having the outcome of interest does not vary with dose (i.e., that the slope has value zero) and the one-sided alternative hypothesis that the probability increases with increasing dose (i.e., that the slope is greater than zero). Analogous approaches were taken for the other response variables (TSH, etc.). See section VIII.C.2 below for more details.

One problem that arises in using a linear probability model is the following: if the slope is greater than 0 (or less than 0), then for sufficiently large doses the model will yield probabilities greater than 1 (or less than 0), which are not permissible values. This could present a practical problem for disease outcomes with low background rates, since slightly negative slopes might imply disease probabilities less than 0 for doses with the range that occurred among study participants. As discussed further below, other models for the dose-response relationship were examined as alternatives to the linear model. One of these, the logistic model (described in section VIII.C.2), has the practical advantage that the probabilities derived from it are always greater than 0 and less than 1, regardless of the values of the intercepts, regression coefficient or dose. Therefore the logistic model was employed not only as an alternative to the primary linear model, but also for more detailed investigation of the influences of other factors on the radiation dose-response.

A.1.b. Confounding and Effect Modification

The relationship between disease risk and a possible risk factor such as radiation exposure is said to be "confounded" if both the risk of disease and exposure are correlated with some other factor, called a confounding factor or simply a confounder. If the presence of confounding is ignored, an epidemiological study can produce erroneous results. Suppose, for purposes of illustration, that smokers received higher doses from Hanford's ^{131}I than nonsmokers, and that smoking itself increases the risk of thyroid disease.

Finally suppose, also for this example, that disease risk is unrelated to radiation exposure. Then if a study simply examined the relationship between thyroid disease and radiation dose without accounting for smoking, it might erroneously conclude that disease risk is higher among people more heavily exposed to radiation compared to those with less exposure, because the former group included more smokers. Confounding can also cause a study to conclude erroneously that there is no association between an outcome and an exposure that in fact increases risk of the outcome. The potential problem of confounding can be addressed in epidemiological studies by performing analyses that adjust for the effects of possible confounders.

Effect modification occurs when the association between disease risk and the exposure of interest differs according to a third factor, called the "effect modifier." For example, an association between risk of a certain thyroid disease outcome and radiation exposure might occur only among women, but not among men. In that situation, sex would modify the radiation effect.

Identification and analysis of confounding and effect modifying factors was accomplished through the analysis of generalizations of the logistic dose-response models mentioned above. For disease outcomes, these generalizations allowed the background probabilities of the outcome of interest (i.e., the intercept parameters) and/or the regression parameters to vary as functions of factors in the following categories:

- Sex
- Age at first exposure to ^{131}I from Hanford
- Age at HTDS examination
- Ethnicity
- Smoking
- Other radiation exposure to the thyroid (occupational, medical, dental, fallout from the Nevada Test Site)

In addition, the source of each participant's dosimetry data, i.e., the CATI or the Expanded In-Person Interview, was included among the potential confounding or effect modifying factors.

A.1.c. Shape of Dose-Response Relationships

Investigation of the shapes of dose-response relationships was accomplished through the analysis of generalizations of and alternatives to the primary linear dose-response model, including linear-quadratic and logistic models.

A.2. Estimation and Significance Testing

In drawing inferences about dose-response relationships, two general statistical approaches may be considered: estimation and significance testing. These two approaches are largely complementary, each providing useful information that the other does not, and each is needed to meet the study's overall objective. Therefore both approaches were employed in reporting results of the HTDS. Regarding Objective 1, for example, regression coefficients that represented how each response variable listed in section VIII.A.1.a above changes in relationship to the ^{131}I radiation dose to the thyroid were estimated. These estimates included confidence intervals, which serve to characterize how precisely the true values of the coefficients were likely to have been estimated. In addition, however, significance tests were performed. The one-sided p-values produced by these tests indicated the degree to which the study results were inconsistent with, and therefore evidence against, the null hypotheses that the outcomes are not associated with dose. Thus the two approaches together provided estimates of the magnitude of any radiation effects, and measures of the strength of evidence against the null hypotheses of no association.

B. Definitions of Variables

The kinds of data that were collected and the analyses that were performed for the study can be divided into three categories:

1. Process information. This includes descriptive analyses regarding the numbers of persons selected, the success rates of the various steps in locating and recruiting those persons, and in completion of the study's various data collection activities.
2. Characteristics of living evaluable participants. This includes descriptive analyses regarding demographic variables, as well as characteristics used for the calculation of dose estimates, and information about possible occupational and medical exposures to radiation.
3. Analyses of exposures and outcomes. This includes descriptive analyses of the distributions of radiation dose to the thyroid and of frequencies of disease outcomes and other response variables, as well as the inferential analyses of the radiation dose-response relationships, including analyses of dose effect modification and confounding, and of the effect of uncertainty in the dose estimates.

B.1. Process Information

HTDS used computerized tracing and tracking systems to monitor the progress of the 5199 selected persons through the study's various steps of identification and location (tracing), and contacting, recruitment, and study participation (tracking). At the end of data collection, these systems contained information about the outcomes of the various steps for each selected person. This information was used to describe the success rates of the various steps, and to search for possible sources of bias that might affect the estimated dose-response relationships. The following variables were obtained from the tracing and tracking data:

B.1.a. Stratification Factors

Stratification factors included sex, year of birth, and mother's usual place of residence ("geostratum"), as recorded on the selected person's birth certificate. From the tracing, recruiting and interviewing activities of this study, it was noted that the sex or birth year data were incorrect for a small number of selected persons. The resulting corrections were not made in the stratification data (since there was less or no possibility of detecting such errors for persons who were not located or not interviewed); corrected sex and year of birth data were recorded in separate data files. All analyses involving sex and birth year in this report are based on the corrected data, unless specifically indicated otherwise.

B.1.b. Tracing Outcome

At the end of the tracing component of the study, each of the 5199 selected persons was categorized as not located; located, deceased; or located, alive.

B.1.c. Cause of Death

For all selected persons who were found to be deceased when located, or who were located alive but died prior to meeting the criteria that define a living evaluable participant, death certificates were sought, and the causes of death abstracted. The causes were categorized, taking into account the need to identify conditions related to thyroid or parathyroid disease. In addition, the primary cause of death for each deceased cohort member for whom a cause of death was identified was coded using the International Classification of Diseases, 9th Revision (ICD9-CM). For deaths prior to 1979, (when ICD9-CM was implemented) the primary cause of death was also coded to the system in use at that time.

B.1.d. Contacting Outcomes

At the end of the recruiting component of the study, each person who was located alive was categorized as to whether or not he or she could be contacted by telephone call from the HTDS recruiting staff.

B.1.e. Recruiting Outcomes

At the end of the recruiting component of the study, each contacted person was categorized according to the outcome of the recruiting effort, as either agreed to participate, refused to participate, or lost to contact without agreeing or refusing. Those who agreed or refused were also classified according to whether they agreed or refused on the initial recruiting attempt or after recontacting. Those who failed to decide whether or not to participate by the end of the recruiting period or who initially agreed to participate but subsequently withdrew that agreement were counted as having refused.

B.1.f. Dosimetry Data Collection

Data for dose estimation were collected from two sources. The preferred source of data was the Computer Assisted Telephone Interview (CATI); (see section V.D above, and HTDS Protocol, Appendix 1) of one or more persons with direct knowledge of the participant's infancy and childhood and, for participants born after mid-December, 1994, the participant's mother's pregnancy. In some cases, however, no suitable and willing CATI respondent could be identified, or the information from the CATI was judged to be unreliable. In such cases, the dosimetry information was collected from the participant by means of an expanded version of the In-Person Interview (Exp-IPI) conducted during his or her clinic visit. At the end of the CATI and clinic components of the study, each person who agreed to participate was categorized according to: 1) whether or not a CATI was completed; and 2) whether or not an Exp-IPI was performed. Persons for whom a CATI was performed were also categorized according to the relationship of the primary CATI respondent to the subject of the interview: birth mother, adoptive mother, father, sister, brother, aunt, uncle, other relative, or other. In some instances, CATI data were collected but not used for dose estimation, or CATI and Exp-IPI data were combined (e.g., for participants for whom the CATI was judged inadequate). For purposes of analysis, the participants were classified according to the source of their dosimetry data: CATI versus Exp-IPI.

The CATI included information about the Interviewer's assessment of quality of responses. This information was collected at several points during the interview: following sections concerning sources of milk, the mother's milk consumption and dietary history (if applicable, i.e., if the participant was born after December 15, 1944), the participant's milk consumption and dietary history, the mother's medical history, and the participant's medical history; and after completion of the entire interview. At each of these points the Interviewer recorded his or her subjective assessment of the quality of the responses (high, generally reliable, questionable, or unreliable). If the quality was rated unreliable or questionable, the Interviewer also recorded his or her subjective assessment of the main reason (unclear memory of events, uncertain understanding of questions, hurried responses, or other). In addition, following the sections concerning the participant's milk consumption and dietary history and the participant's medical history, the Interviewer recorded his or her subjective assessment of how often explanatory text was repeated (very often, often, not often). At the end of the CATI the Interviewer recorded his or her subjective assessment of the respondent's cooperation (very good, good, fair, poor).

The Exp-IPI included much more limited information about the Interviewer's assessment of interview quality. At the end of the interview, the Interviewer recorded his or her subjective assessments of the respondent's cooperation and the quality of responses (using the categories defined above). In addition, for interviews rated unreliable or questionable, the Interviewer recorded his or her subjective assessment of the main reason for this rating in narrative form. These narrative answers were classified into the categories defined above for the CATI.

B.1.g. Clinic Participation

Each person who agreed to participate was categorized according to whether or not he or she attended a clinic. Participants (i.e., persons who attended a clinic) were also classified according to whether each of the clinical components was completed: In-Person Interview, ultrasound examination, radiologist review of ultrasound examination, blood draw, thyroid function tests, and physical examination of the thyroid. For participants who received the physical examination of the thyroid, the number of examining physicians, one or two, was recorded. In addition, for participants recommended to have a fine needle aspiration, thyroid scan, or other follow-up for diagnosis of thyroid or parathyroid disease, the results of those procedures were also recorded.

B.1.h. Requests for Medical Records or Slides

Each request for medical records or slides was classified according to the type of request: past medical records, past pathology slides, post-clinic medical records, and post-clinic pathology slides. The "post-clinic" requests refer to records or slides that were created after the participant's clinic visit and as a result of an HTDS recommendation for further evaluation. In addition, each request was classified according to outcome: requested materials received versus not received. For each living evaluable participant the number of requests of each of the four types was recorded, along with the corresponding numbers of requests for which materials were received.

B.2. Characteristics of Living Evaluable Participants

B.2.a. Demographic Data

The following demographic variables were obtained from the tracking system and interview results: sex (corrected), year of birth (corrected), age at HTDS examination, race/ethnicity, religious preference.

B.2.b. Residence History

A residence history is a description of the places a person has lived, and of the dates he or she lived at each place. Residence histories for study participants ranged from the very simple (e.g., a single residence throughout the entire period) to the very complex (e.g., dozens of residences during the period). For each living evaluable participant, the number of residences in the HEDR domain and the duration of residence in the HEDR domain during the period of interest (December 1944 through December 1957) were determined from the CATI or Exp-IPI as appropriate. Note that some living evaluable participants who were born and moved away from the domain before December 15, 1944 had no residences within the HEDR geographic domain during the time period of interest. These participants, designated out-of-area participants, were not excluded from the study.

B.2.c. Dosimetric Data

Dosimetric data includes the information (other than residence history) that is used to calculate an individual's estimated dose, such as the consumption levels and sources of milk and food products. Most of the data used for calculating dose estimates has the characteristic of varying over time. Key determinants of the radiation dose to the thyroid, such as sources of food products and quantities consumed, are subject to change at unpredictable points in time and cannot be characterized by single numerical or categorical variables.

For participants with CATIs as the source of dosimetry data, the participant's consumption levels of the following milk and food products were recorded: processed cow's milk, raw cow's milk, the total of processed and raw cow's milk, processed goat's milk, raw goat's milk, the total of processed and raw goat's milk, fresh fruit, fresh green and leafy vegetables, and eggs from free range chickens. The milk consumption values were reported in the units of grams and 8 ounce servings per day; fruit and vegetable consumption were reported in grams per day; and egg consumption was recorded in grams per days and eggs per week (0, 1, 2, 3, 4-6, 7, > 7).

For participants born or breast-fed after December 14, 1944, the CATI included questions about the mother's sources and consumption of milk and other food products. These were used to record the mother's consumption levels of the following for the period from December 15, 1944 until the participant's birth or end of breast-feeding as appropriate: processed cow's milk, raw cow's milk, the total of processed and raw cow's milk, processed goat's milk, raw goat's milk, the total of processed and raw goat's milk, fresh fruit, fresh green and leafy vegetables, and eggs from free range chickens.

For descriptive purposes, consumption data for milk and food products from the CATI was summarized in two ways. The first way was used to show how consumption levels changed with age: each participant's consumption of a particular milk or food product was reported for any of the following dates that fell within the period of interest (December, 1944 through December, 1957): the six-month anniversary of the participant's birth, and each of his or her first through 15th birthdays.

The second summary of consumption data was used to examine how overall milk and milk product consumption levels were correlated with estimated thyroid radiation dose. To calculate each participant's average consumption level, his or her reported total number of 8 oz. servings for a particular type of milk was first calculated by integrating the reported consumption levels over the time periods for which the CATI respondent reported consumption levels of that milk. For example, if a CATI respondent reported that a participant consumed three 8 oz. servings per day over a period of 2 years, the total consumption was $3 \times 2 \times 365 = 2190$ 8 oz. servings for that period. For these calculations participants born in 1946 were assigned milk consumption values of 0 for 1945. Also participants in the 1940-1945 birth strata who never lived inside the HEDR domain during 1945 were assigned consumption levels of 0 for 1945. If the consumption level for a particular type of milk was unknown for any or all of the time period in question (because the CATI respondent could not report the quantity of glasses consumed), the total consumption was considered unknown for these calculations. Two measures of average consumption were calculated. The first, designated "Average No. of 8 oz. servings per day," was obtained by dividing the reported total number of 8 oz. servings for a given time period by the duration of that period in days (e.g., by 365 for average consumption during 1945). The second measure of average consumption, designated "Average No. of 8 oz. servings per in-area day," used a different divisor: the number of days during the period for which (1) the participant lived within the HEDR domain and (2) the level of milk consumption was reported in the participant's CATI. Average consumption levels were calculated for two time periods: (1) 1945, the year in which by far the largest amount of ^{131}I was released from Hanford (see section IV.A.2 above), and (2) the entire period 1944-1957.

B.2.d. Age at Exposure

Age at exposure to ^{131}I may be a particularly important effect-modifying factor: exposure at younger ages may produce a greater increase in risk of subsequent thyroid neoplasia, and perhaps of other outcomes, compared to exposure at older ages. An assumption of such age dependence is built into the NCRP risk estimates for thyroid carcinogenesis induced by exposure to radioiodine (36). However this assumption relies heavily on extrapolation from human studies of other kinds of radiation exposure, and on animal studies. Therefore particular attention was paid to analyzing the effect of age at exposure. Unfortunately, age at exposure was not simply defined for this study, since most participants' exposures to ^{131}I from Hanford occurred over a protracted period of time, and therefore over a range of ages. Therefore age at <u>first</u> exposure was examined as the possible effect-modifying factor.

For calculating age at first exposure to ^{131}I from Hanford, the definition of age was generalized from its usual definition to include negative values representing gestational ages, extending from birth back to gestational age 90 days (about −0.5 years), the age at which thyroid function is assumed to begin in the HEDR model (126). Similarly, the definition of a participant's residence was generalized to include the participant's mother's residence during the participant's gestation from age −0.75 to 0 (birth). With these conventions, age at first exposure was defined as the maximum of −0.75, age on December 15, 1944, and age when the participant first resided in the HEDR geographical domain.

B.2.e. Medical and Dental Radiation Exposure History of Participant

For participants with CATIs as the source of dosimetry data, information about medical and dental radiation exposures was obtained by combining data from the CATI and In-Person Interviews; otherwise the information was obtained from the In-Person Interview alone. For descriptive purposes, each living evaluable participant was classified according to whether or not he or she had a history of each of a number of diagnostic radiation procedures: CAT scan of the upper body, diagnostic x-ray of the head, diagnostic x-ray of the neck, diagnostic x-ray of the chest or upper body (including mammograms), diagnostic x-ray of the stomach or mid-back, barium enema, upper GI, intravenous pyelogram, fluoroscopy of the upper body, thyroid nuclear scan, and other nuclear scan. Participants were also classified according to whether they had a history of the following types of radiation treatment: radiation treatment for any cancer other than thyroid cancer, x-ray treatment to the upper body for acne, x-ray treatment for ringworm, x-ray treatment for enlarged tonsils, x-ray treatment to the upper body for tuberculosis, x-ray treatment for scalp infection, x-ray treatment for enlarged thymus, and x-ray treatment to the upper body for any other reason. Finally, participants were classified according to whether they ever had routine dental x-rays, ever had routine dental x-rays more than once per year, and ever had dental x-rays that did not usually include shielding of the neck area.

B.2.f. Occupational History

The In-Person Interview included questions about employment in a number of industries or occupations that might involve exposure to ionizing radiation. For descriptive purposes, each living evaluable participant was classified according to whether or not he or she had ever worked in each of the following industries and occupations: geology; metallurgy; metal processing; ore refining; mining; nuclear industry; on the premises of a nuclear facility; health care with exposure to radioactive materials or x-rays; scientist, researcher or student with exposure to radioactive materials or x-rays; military working around nuclear testing, nuclear submarines or other radiation exposure; any other industry or occupation that might have caused exposure to radioactive materials or x-rays.

B.2.g. Smoking History

Information about smoking histories was obtained from the In-Person Interview. Participants were categorized according to history of ever smoking each of filtered cigarettes, nonfiltered cigarettes, any cigarettes, cigars, or pipe. In addition, for those who reported ever smoking a particular product, the level of use of that product was quantified in terms of cigarette pack-years (average number of 20-cigarette packs per day times number of years cigarettes smoked), cigar-years (average number of cigars per day times number of years cigars smoked), and pipe-years (average number of bowls per day times number of years pipes smoked) as appropriate.

B.2.h. Exposure to ^{131}I from the Nevada Test Site

Information released by the U.S. National Cancer Institute (NCI) in 1997 (131), indicated that persons living in the contiguous 48 states during the 1950s and 1960s were exposed to various levels of ^{131}I released from the Nevada Test Site (NTS). The information released by NCI included estimates of dose for representative individuals in all counties in the 48 states, as well as more detailed data regarding estimated dose by shot (i.e., by individual test detonation), county, and age. Limited preliminary comparisons for HTDS participants suggested that in many cases the reported NTS dose estimates were comparable to or even greater than the estimated Hanford doses. Therefore it was judged necessary to add exposure to ^{131}I from the NTS to the list of potential confounding factors.

For HTDS, the "estimated NTS dose" was defined as the thyroid dose from ^{131}I entering the atmosphere from tests conducted at NTS between 1951 and 1957, inclusive, as estimated from data made publicly available by NCI. The limitation to tests conducted through 1957 was based on two considerations: 1) although NCI reported exposures through 1972, it was estimated that 99% of the ^{131}I was released from 90 tests conducted between 1952 and 1957, and 2) HTDS collected complete residence histories for all living evaluable participants only through 1957, including residences outside the HEDR domain. Each living evaluable participant's estimated NTS dose was calculated as the total of doses from all 57 shots at the NTS between 1951 and 1957. HTDS staff wrote computer code to accumulate for each participant, the estimated thyroid dose from each shot taking into account the participant's residence history.

B.3. Analyses of Exposures and Outcomes

B.3.a. Exposure Data

The primary analyses of dose-response relationships were based on individual estimates of radiation dose to the thyroid, specifically organ doses to the thyroid which were estimated from the residence history and dosimetric data collected during the CATI and/or Exp-IPI. The participants were divided into two categories regarding dose estimates:

- The first category, and by far the largest, includes the participants who lived at some time between December 15, 1944, and December 31, 1957, in the geographical domain defined by the Hanford Environmental Dose Reconstruction Project. Doses for these participants were calculated using the CIDER program, which was created by the HEDR Project. These are designated **in-area participants**.

- The second category consists of persons who never resided within the HEDR domain between December 15, 1944, and December 31, 1957. The CIDER program does not provide dose estimates for these participants. These are designated **out-of-area participants**.

The dose estimates produced by CIDER for the in-area participants were derived from information collected during the CATI and/or Exp-IPI. After review and editing, these data were formatted into scenario files that served as input to the CIDER program (132). The CIDER output for each in-area participant consisted of 100 realizations of the estimated cumulative total organ dose to the thyroid from ^{131}I, as well as corresponding sets of realizations of dose by year and by pathway. The CATI and Exp-IPI also included a short series of questions meant to elicit the respondent's level of knowledge and opinions regarding thyroid disease, radiation, and Hanford, which were used to investigate the possibility of recall bias.

It is important to recognize that in the CIDER program each of the 100 realizations of dose is calculated for a fixed set of conditions regarding the source term and environmental transport, and that these conditions for a given realization were the same for every participant. The 100 realizations were

obtained by randomly varying the conditions, i.e., the uncertain parameters in the HEDR models for source term, transport, etc., in order to characterize the uncertainty in the resulting dose estimates (16). Thus it is useful to view each realization as consisting of a set of doses, one for each in-area participant. This can be illustrated by the following table, in which the k-th realization of dose for the i-th participant (i = 1, …, N) is denoted $D_{i,k}$, where N is the number of living evaluable in-area participants.

Table VIII.B-1. Schematic Illustration of Dose Realizations

	Realization			
Participant	1	2	…	100
1	$D_{1,1}$	$D_{1,2}$	…	$D_{1,100}$
2	$D_{2,1}$	$D_{2,2}$	…	$D_{2,100}$
…	…	…	…	…
N	$D_{N,1}$	$D_{N,2}$	…	$D_{N,100}$

Each column in this table, i.e., each realization $\{D_{1,k}, \ldots, D_{N,k}\}$, is a set of doses which are consistent in the sense that they were all calculated under the same conditions. For example, the amounts of ^{131}I released into the air (the "source term") will be higher in some realizations and lower in others. This variability is likely to induce a corresponding variation in dose estimates: realizations with higher or lower source terms may tend to produce higher or lower dose estimates, respectively, for many participants. As a result, the dose estimates of different participants may tend to be correlated across the 100 realizations. Some components of the dosimetry model, for example those subject to the constraint of mass balancing, may introduce negative correlations. Consider the example of atmospheric transport. For each realization to be properly mass-balanced, if one region receives a particularly high deposition of ^{131}I, then the depositions in other regions may tend to be lower. Thus estimated doses of participants exposed largely from the depositions in the first region may tend to be negatively correlated with the doses of those exposed to ^{131}I deposited in other regions.

In the original version of the CIDER program, the dose conversion factors (DCFs), which in effect convert estimated amounts of ^{131}I taken up by the thyroid (measured in Ci) into estimated dose (in mGy or related units), were assumed to be the same for all participants in each realization. This almost certainly induced an unrealistically high level of positive correlation: every participants' dose estimates would tend to rise or fall together as the DCFs increased or decreased from realization to realization. Therefore the CIDER program was modified to permit the realizations of DCFs to be randomly permuted for each participant (see Appendix 22). This was expected to greatly reduce the correlation of dose realizations across participants.

A further revision of the CIDER program allowed uncertainties to be applied to dietary input data for CATI participants. Incorporating this additional source of uncertainty would of course increase the uncertainties of the resulting dose estimates. Since the magnitudes of these uncertainties could not be determined or estimated from the data collected for HTDS or from other sources, the revision of CIDER allowed their magnitudes to be specified (see Appendix 22). As described further below, this capability was used to assess how estimates of radiation dose-response parameters were affected by the incorporation of additional uncertainties of various plausible magnitudes.

For many purposes it was useful to have a single number or "point estimate" to represent each participant's dose. For each in-area participant, the median of the 100 realizations of dose, $d_i = \text{median}(D_{i,1}, \ldots, D_{i,100})$ for participant i, was used as a summary measure of that participant's dose. In particular, the median doses were used for descriptive purposes that required categorization of participants by dose. Two other point estimates were also calculated for each in-area participant. The first is the geometric mean:

$$GM_i = \exp(N^{-1} \Sigma_k \ln D_{i,k}).$$

Finally, for comparability with reported results of the Utah Thyroid Study (10479, 11425, 10139), the arithmetic mean dose (called simply the mean dose) was also calculated:

$$M_i = N^{-1} \Sigma_k D_{i,k}.$$

For descriptive purposes, it as also useful to have summary measures of the uncertainty of dose estimates for the in-area participants. For use with the median doses, the ratio of the 95th percentile to the median dose was calculated. In addition, the geometric standard deviation (GSD_i) and standard deviation (SD_i) were calculated for use with the geometric and arithmetic means, respectively:

$$GSD_i = \exp([(N-1)^{-1} \Sigma_k (\ln^2 D_{i,k} - N^{-1} [\Sigma_k \ln D_{i,k}]^2)]^{1/2})$$

$$SDi = [(N-1)^{-1} \Sigma_k (D_{i,k}^2 - N^{-1} [\Sigma_k D_{i,k}]^2)]^{1/2}.$$

Preliminary analysis of doses indicated that the empirical distributions of the logarithms of the individual participant's doses, ln(Di,1), … , ln(Di,100), were roughly normally distributed, with variances that changed relatively little from participant to participant. As a result each participant's median and geometric mean were nearly equal. Thus analyses of dose-response relationships were essentially unchanged whether based on medians or geometric means. Also, the arithmetic mean doses were roughly a constant multiple, C, of the median or geometric mean where C > 1. Preliminary results suggested that the value of C will be about 1.35, i.e., that the mean doses were about 35% larger than the median or geometric mean doses. Since mean doses were expected to be consistently larger than geometric mean or median doses, the estimated effects of exposure on outcomes were expected to be smaller in magnitude if based on mean dose, compared to median or geometric mean dose.

B.3.b. Alternative Representations of Exposure

In addition to the individual estimates of thyroid radiation dose described above, alternative representations of exposure to Hanford's ^{131}I were defined. When the HTDS protocol was developed, the consideration of such alternatives arose from the possibility that the HEDR project, which had not then completed its feasibility phase, might not provide a system for calculating individual dose estimates. Since the HTDS did in fact develop a dosimetry system that could be adapted for HTDS use, this reason for considering alternative characterizations of exposure became moot. Nevertheless, alternative characterizations remained of interest, since they could be used to assess whether there might be evidence of a radiation effect that was not revealed in the primary dose-response analyses using individual dose estimates from the CIDER program. Two alternative representations of exposure were considered: geostratum and a dichotomous (high versus low) exposure variable. Unlike the estimates of thyroid radiation doses, which were available only for the in-area participants, both of these alternative representations of exposure were defined for all living evaluable participants, including the out-of-area group.

B.3.b.1 Geostrata

The first alternative was simply the participants' geostrata, i.e., the nine geographical regions that were defined for the selection of potential study participants (see section IV.A.1 above). The rationale for considering geostratum as an alternative representation of exposure was as follows. The results of the HEDR project strongly suggested that doses received by participants varied markedly according to their places of residence, particularly during the period of highest ^{131}I releases (130). Since each participant's geostratum was his or her mother's usual place of residence at the time of the participant's birth, many participants were likely to have resided in their respective geostrata for at least some of their infancy or

childhood. Therefore geostratum might be at least somewhat correlated with the doses study participants received.

There are obvious limitations in using geostratum as an alternative representation of exposure. Most importantly it fails to account for changes in residence or dietary factors that can strongly influence the dose an individual actually received. Therefore analyses of cumulative incidence of disease outcomes or prevalence of thyroid UDAs in relation to geostrata were unlikely to provide conclusive evidence either for or against an effect of ^{131}I from Hanford.

B.3.b.2. Dichotomous Exposure Variable

The second alternative was defined in such a way as to reduce the weaknesses inherent in using geostratum as a characterization of exposure. Specifically, an attempt was made to assign the living evaluable participants into relatively high and low exposure groups using simple characterizations of the residence and milk consumption histories. The high exposure group was defined to include participants who lived in the downwind counties closest to Hanford during 1945 and consumed appreciable quantities of milk, while the low exposure group was defined to include participants who lived sufficiently far away from Hanford and/or drank sufficiently small quantities of milk. Specifically, the two groups were defined as follows:

- High exposure group. This group included all living evaluable participants born before July 2, 1945, who lived in Benton (excluding the city of Richland, but including Kennewick), Franklin (including Pasco), or Adams County for at least 180 days during 1945, and who were reported to consume an average of at least one 8 oz. serving of milk and milk products per day during 1945. Since this criterion depends in part on the participant's individual milk consumption history, only participants with CATI data used for dose estimation could be included in this group.

- Low exposure group. This group included living evaluable participants in the following categories:

  (i) Out-of area participants.

  (ii) In-area participants born before January 1, 1946 who lived in Ferry, Stevens, or Okanogan County or outside the HEDR domain from December 15, 1944 or their birthdays (whichever occurred first) through December 31, 1951.

  (iii) In-area participants born before January 1, 1946 who lived in Ferry, Stevens, or Okanogan County or outside the HEDR domain from December 15, 1944 or their birthdays (whichever occurred first) through December 31, 1945, and who lived outside Benton, Franklin and Adams Counties from January 1, 1946 through December 31, 1951.

  (iv) In-area participants born before January 1, 1946 who are not in categories (i) or (ii), but who lived outside Benton, Franklin and Adams Counties from December 15, 1944 or their birthdays (whichever occurred first) through December 31, 1951, and who were reported to consume an average of less than one 8 oz. serving of milk and milk products per day during 1945. Note that only participants with CATI data used for dose estimation could meet this criterion.

  (v) In-area participants born after December 31, 1945 who lived outside Benton, Franklin and Adams Counties from birth through December 31, 1951.

While the exposure groups defined above were expected to provide a more reliable characterization of exposure than geostratum, they could not be expected to provide a perfectly accurate separation of high- and low-exposed participants. For example, it could not be assured that every

participant in the high exposure group had a higher dose from Hanford's ^{131}I than every participant in the low exposure group. The criteria above were defined to ensure a reasonable likelihood that the high and low exposure groups consisted largely of participants with comparatively high and low doses, respectively.

The categories defined above did not include all possible circumstances, e.g., participants who lived at least 180 days in Benton, Franklin or Adams Counties during 1945 but who were reported to consume an average of less than one 8 oz. serving of milk and milk products per day during 1945. Living evaluable participants who did not meet any of the criteria above for either the high or low exposure group were not assigned to either group, and were excluded from analyses involving the dichotomous exposure variable.

B.3.c. Outcome Data

The outcome data for this study included the following:

- Diagnoses of thyroid disease and primary hyperparathyroidism
- Presence of ultrasound-detected abnormalities of the thyroid (UDAs)
- Results of thyroid function and antibody tests, thyroid volume, and serum calcium levels

See section VIII.A.1 a. above for a more detailed list of outcomes, and IV.C for definitions of outcomes. Data for the first two categories (diagnoses of thyroid disease and hyperparathyroidism, and ultrasound-detected abnormalities) were obtained from the Final Diagnosis Determination Form (FDDF; Appendix 20). Thyroid mass was obtained from the ultrasound measurements recorded by the HTDS sonographer at the time of the clinical examination (LxWxHx.55 calculated separately for the right and left lobes, then added together), unless revised by the radiologist. Results of thyroid function and antibody tests and serum calcium levels were obtained from reports provided by the laboratories performing the analyses.

One disease outcome category from the FDDF requires special comment: "thyroid nodule suspicious for follicular neoplasm." This category was included on the FDDF to allow for the possibility that diagnostic information would not permit a definitive determination of a nodule's behavior (benign versus malignant). It was also used for participants who, on HTDS cytology review, were described as having a nodule with an intermediate or high probability of being a follicular neoplasm, and who did not have a subsequent surgery that could provide a definitive histologic diagnosis. At the end of the data collection period, all participants remaining in the "suspicious" category were in this latter group. None of these cases were suspicious for papillary carcinoma. Most of the nodules in this group were likely to be benign, however they could not be counted among the confirmed cases of benign nodules. Therefore they were included in the combined category of participants with any thyroid nodule. In addition, they were included along with the benign nodules in a secondary analysis to assess whether their omission might disguise a dose-response.

For each participant with a disease outcome, additional information about that outcome was available from the FDDF, including the basis for the diagnosis and, for some outcomes, possible etiologies or contributing causes (see Appendix 20). Similarly, for participants with a thyroid UDA, the FDDF included further information about the UDA. Therefore it was important to identify a primary definition for each outcome, as well as alternative definitions of outcome that would be considered. The primary definitions were intended to include cases with (1) a broad but meaningful range of specific outcome subtypes (e.g., benign thyroid nodules of any histologic/cytologic type), and (2) an adequately definitive basis for diagnosis (e.g., based on histologic or cytologic evidence confirmed by the HTDS evaluation). The alternative definitions were intended to permit analysis of the effects of (1) restricting outcomes to more specific subtypes (e.g., benign thyroid nodule excluding nonneoplastic disease, or non-iatrogenic hyperthyroidism), and (2) changing the level of diagnostic certainty (e.g., including all diagnoses, ranging

from those based on HTDS evaluation to those based only on participant/respondent report). The primary and alternative definitions of the various disease and ultrasound outcomes are given in section IV.C.

C. Analytic Methods

A variety of descriptive analyses were performed to summarize process information, characteristics of the living evaluable participants, and exposure and outcome data. These analyses made use of standard descriptive statistical techniques, primarily frequency tables and crosstabulations, calculation of estimates of central tendency (e.g., means, medians) and dispersion (e.g., ranges, 5^{th} and 95^{th} percentiles), and simple plots of standard types (e.g., bar plots, pie charts, scatter plots).

C.1. Statistical Models for Analyses of Exposures and Outcomes

Standard statistical techniques were used to provide descriptions of radiation dose estimates and outcomes. These included the calculation of summary statistics (median, minimum, maximum, mean, standard deviation) and cumulative distribution functions to describe distributions of estimated doses, for presentation in tabular or graphical form. Uncertainties of dose estimates for in-area participants were illustrated graphically, including cumulative distribution functions of ratios of the 95th percentile to the median dose, and of geometric standard deviations (GSD) of dose, and scatterplots of the 95th percentile-to-median dose ratio (on the vertical axis) by median dose (horizontal axis, on logarithmic scale), and of GSD by geometric mean dose.

Summaries of outcome data were displayed in tables showing the numbers of cases and relative frequencies for subcategories defined by, e.g., basis for diagnosis or disease subtype, for women and men separately, and for both sexes combined.

The relationships between outcomes and estimated dose were displayed in tables showing numbers of cases within dose categories, for women and men separately and for both sexes combined. The number of cases in each category was also expressed as a percentage of the number of living evaluable participants in the category.

C.1.a. Inferences About Dose-response Relationships: Models for Objective 1

As described in the HTDS Protocol, Appendix H (1), the primary analysis of exposure-outcome relationships for disease outcomes focused on the cumulative incidence of the outcome among living evaluable participants at the time they are examined for the study. "Cumulative incidence" referred specifically to the proportion of participants with the outcome of interest diagnosed at any time up to and including the HTDS examination. Thus it is most comparable to "period prevalence" as defined for the Utah Thyroid Study (133), if it is understood that the beginning of the period of observation is the birth of the participant. However the term "cumulative incidence" was used for HTDS, since "period prevalence" implies a risk period defined by uniform calendar dates for all study participants. One basic model served as the starting point for estimation and significance testing of the dose-response relationships for the disease and ultrasound outcomes listed in paragraph II.B above. This was the stratified linear probability model:

[1] $P_j(d) = A_j + B \times d$

where

j = 1, 2 indexes the strata defined by sex,
d is the cumulative dose to the thyroid,

$P_j(d)$ is the probability that a living evaluable participant in stratum j and with dose d has the disease of interest,
A_j is the background probability for participants in stratum j, i.e., the probability of the outcome in the absence of the radiation exposure, and
B is the regression coefficient that expresses the magnitude of the radiation effect.

Like disease outcomes, the presence or absence of thyroid UDA is also a binary outcome, and model [1] is applicable. However unlike the disease outcomes, for which diagnoses could have occurred any time up to and including the HTDS examination, detection of thyroid UDAs was based entirely on the HTDS examination. Therefore for thyroid UDAs, the probability $P_j(d)$ in model [1] refers to the prevalence, rather than cumulative incidence.

The regression coefficient B in [1] represents the slope of the dose-response. According to the model, the probability of disease increases with increasing dose, does not change with dose, or decreases with increasing dose depending on whether $B > 0$, $B = 0$, or $B < 0$, respectively. Suppose, for example, that the background probability of thyroid cancer among women is $A_1 = 0.007$. Table VIII.C-1 illustrates how the probability of thyroid cancer varies in relation to dose for three different values of the slope B.

Table VIII.C-1. Illustration of Positive, Zero, and Negative Dose-responses

	Probability of Thyroid Cancer for Women ($A_1 = 0.007$)		
Dose (mGy)	B = 0.025 per Gy	B = 0.000 per Gy	B = –0.005 per Gy
0	0.0070	0.0070	0.0070
100	0.0095	0.0070	0.0065
1000	0.0320	0.0070	0.0020

Note that if $B < 0$, then for sufficiently large doses, the probability of disease will be less than or equal to 0, especially for outcomes with low background rates. For example, continuing the illustration from the table above, if B = –0.005 per Gy, then for women with doses greater than 1400 mGy (1.4 Gy) the linear probability model implies that the probability of disease is less than 0, which is impossible. While it is almost certain that probabilities of disease outcomes do not decrease with increasing dose, the estimate of B may be less than 0 due to the essentially random variability of disease occurrence, especially if the true value of B is near 0. Therefore the parameters of model [1] were estimated under the constraint that every participant's probability must be greater than 0. (Estimation was similarly constrained to ensure that every participant's probability is less than 1, the maximum possible value for probabilities. However this constraint was rarely invoked since most outcomes had sufficiently low background rates.)

Since the linear probability model could yield impermissible values for probabilities, e.g., cumulative incidence of disease or thyroid UDA prevalence less than 0 (see section VIII.A.1.a above), the sex-stratified logistic regression model was also considered.

[2] $$P_j(d) = \exp(A_j + B \times d) / [1 + \exp(A_j + B \times d)].$$

It should be noted that the parameters of the logistic model do not correspond directly to those of the linear model [1]. For example, the values of the intercept parameters A_j are not the background probabilities. Nevertheless the background probabilities can be calculated from the parameters of the logistic model:

[3] $$P_j(0) = \exp(A_j) / [1 + \exp(A_j)] \text{ for } j = 1, 2.$$

Although the regression coefficient B in the logistic model does not represent the slope of a linear probability model, it can nevertheless be used in a similar way to assess the evidence for or against the existence of a dose-response relationship.

A stratified model analogous to [1] was used for laboratory values (thyroid function, antibody tests, and serum calcium) and thyroid volume:

[4] $E_j(d) = A_j + B \times d$

where

$E_j(d)$ is the mean of the (possibly transformed) value for living evaluable participants in stratum j and with dose d, and
A_j is the background mean for participants in stratum j, i.e., the mean in the absence of the radiation exposure, and the other terms are defined as for [1].

Transformation of some laboratory values (e.g., to logarithms) was expected to be appropriate since they are bounded below by zero and likely to be right-skewed. Analyses of TSH, T4, T3RU, and FTI in relation to estimated thyroid radiation dose were limited to participants who were not on thyroid hormone replacement at the time of their HTDS examination.

As mentioned above, three different assays were used for TSH, and two assays for anti-thyroid antibody. The three TSH assays (RIA, EIA-1, and EIA-2) all measured the same quantity (serum concentration of TSH). Therefore for TSH model [4] was generalized to assess whether B and/or the A_j differed among the three assays. The situation was different for the two anti-thyroid antibody assays: AMA and anti-TPO do not measure the same quantity. Therefore their data were not combined, and model [4] was fit separately to the AMA and anti-TPO data.

C.1.a.1. Alternative Point Estimates of Dose to Thyroid from Hanford ^{131}I

As noted above, two point estimates of each in-area participant's dose were available in addition to the median: the arithmetic and geometric means. To assess the extent to which the results might be influenced by the choice of the point estimate to represent dose, certain analyses were repeated using the arithmetic and geometric means.

The use of arithmetic mean doses is similar to the approach taken in the analysis of the Utah Thyroid Study (133, 134). However it must be noted that the results of the HTDS and the Utah study are not directly comparable, even for the analyses of neoplastic diseases, since their outcome variables differ: the HTDS dealt with lifetime cumulative incidence through the early-to-mid 1990s, in a cohort born in the early-to-mid 1940s; the Utah study dealt with incidence and prevalence in the late 1960s and mid 1980s, in a cohort born from the mid-1940s to the mid-1950s.

C.1.a.2. Sensitivity of Results to Large Doses

The distribution of doses was expected to be quite skewed, with large numbers of participants having comparatively low doses, and small numbers having quite high doses. Therefore, for the disease outcomes and thyroid UDAs, analyses were performed to assess whether the regression coefficient B might be inordinately influenced by the high dose participants. In particular, two empirical checks were made to assess whether the estimated regression coefficient adequately represents the dose-response relationship over the lower dose range. The first check consisted of fitting a linear-quadratic exposure-response model:

[5] $P_j(d) = A_j + B_1 \times d + B_2 \times d^2$

If the quadratic term B_2 was found to be significantly different from 0, then the estimated regression coefficient B from the linear model [1] could be interpreted as underestimating or overestimating the effect in the low dose range, depending on whether the estimate of B_2 is negative or positive, respectively. The second check consisted of fitting the linear model [1] with participants in high dose categories excluded.

C.1.a.3. Dose Estimates for Out-of-Area Participants

It cannot be assumed that the out-of-area participants were unexposed to ^{131}I from Hanford. Indeed, results of the HTDS Pilot Study suggested that many out-of-area participants lived in locations near the HEDR domain at various times during 1945-1957 (135). Furthermore, results of the HEDR project strongly imply that people living outside the domain could have received doses higher than those for some people who lived inside the domain; see for example, Figures 6 through 8 of Farris, et al. (10493). The following empirical approaches were taken to provide dose estimates for the out-of-area participants. These dose estimates were used to assess the sensitivity of dose-response results to assumptions about the doses.

- Out-of-area dose assumption 1: All out-of-area participants were assigned doses of 0 mGy.

- Out-of-area dose assumption 2: Each out-of-area participant who lived anywhere in Washington State, Oregon, Idaho, Montana, British Columbia, or Alberta between December 15, 1944, and December 31, 1957, was assigned the maximum dose for a representative child residing in the grid square of the HEDR domain closest to any of the participant's residences. All other out-of-area participants were assigned doses of 0 mGy.

The assignment of maximum doses required for assumption 2 was accomplished as follows. The region lying outside the HEDR domain but within Washington State, Oregon, Idaho, Montana, British Columbia, or Alberta was divided into four subregions, each corresponding to part of the boundary of the HEDR domain.

- North/northeast subregion: British Columbia, Alberta, and counties in the northern halves of Idaho and Montana, corresponding to the northern boundary and upper half of the eastern boundary of the HEDR domain.

- Southeast subregion: The remaining counties of Montana and counties in Idaho lying north of a boundary defined by county lines extending approximately southeast from the southeastern corner of the HEDR domain, corresponding to the lower half of the boundary of the HEDR domain.

- South subregion: The remaining counties in Idaho, and counties in Oregon lying east of a boundary defined by county lines extending approximately southwest from the southwestern corner of the HEDR domain, corresponding to the southern boundary of the HEDR domain.

- West subregion: The remaining counties in Oregon, and all counties in Washington State, corresponding to the western boundary of the HEDR domain.

The maximum estimated dose for a representative child was then calculated for each of four the segments of the HEDR domain boundary to which the subregions correspond. Based on a representative child born in December 1944 with a diet of backyard cow's milk and produce, the associated doses were 51 mGy for the north/northeast subregion, 12 mGy for the southeast subregion, 14 mGy for the south subregion, and 8 mGy for the west subregion.

The residence histories of the out-of-area living evaluable participants were then reviewed to identify those who had ever lived in any of the four subregions between December 1944 and the end of 1957. Those who had were assigned the highest dose for any subregion in which they had lived during that period.

For each disease and thyroid UDA outcome, the sensitivity of the dose-response results to the inclusion or exclusion of the out-of-area participants was assessed by comparing the results from the primary analysis (which excluded the out-of-area participants) to those obtained from the following two scoping analyses. The linear probability model [1] was used for all of these analyses.

- Scoping Analysis #1: Out-of-area participants were assigned doses under assumption 1 (i.e., extremely low doses) if they did not have the outcome of interest, and under assumption 2 (comparatively high doses) if they did have the outcome of interest, i.e., imposing a strong positive dose-response relationship among the out-of-area participants.
-
- Scoping Analysis #2: Out-of-area participants were assigned doses under assumption 1 (i.e., extremely low doses) if they had the outcome of interest, and under assumption 2 (comparatively high doses) if they did not have the outcome of interest. This imposed a strong negative dose-response relationship among the out-of-area participants.

These two scoping analyses were intended to represent a wide but plausible range of impact that the out-of-area participants might have on the estimated dose-outcome relationships.

C.1.b. Inferences About Dose-response Relationships: Models for Objective 2

Generalizations of the logistic regression model [2] were examined to identify and account for confounding and effect-modifying factors in the analyses of disease outcomes and thyroid UDAs. These generalizations permitted the background probabilities to depend on factors in addition to sex (e.g., year of birth, age at HTDS examination, smoking, and other thyroid radiation exposure), and the regression coefficient to depend on those factors as well as sex. To model effects on the background probabilities, A_j was replaced with an expression of the form $a'x$, where x is a vector with components representing the stratification and the additional factors to be considered and a is a vector of corresponding regression coefficients. Similarly, to model effects on the regression coefficient (i.e., to identify effect-modifying factors), the regression coefficient B was replaced by an expression of the form $b'z$, where z is a vector with components representing the factors being considered.

C.1.c. Inferences About Dose-response Relationships: Models for Objective 3

Alternatives to model [1] were considered in order to investigate the shapes of any exposure-response relationships that were found. These included the logistic model [2] and the linear-quadratic model [5] described above.

C.1.d. Inferences About Dose-response Relationships for Numbers of UDAs

Additional analyses of ultrasound-detected thyroid abnormalities (thyroid UDAs) were performed to investigate whether the average number of abnormalities above a given size might increase with increasing radiation dose to the thyroid. For these analyses, each participant's number of such UDAs was assumed to follow a Poisson distribution with mean value $M_j(d)$ for participants in stratum j with dose d, where

$$M_j(d) = \exp(A_j + B \times d).$$

C.2. Calculational Methods for Inferential Analyses

No previous epidemiological studies have dealt with exposure data of the kind available for this study, i.e., correlated sets of multiple realizations of estimated dose for the in-area participants; and no specific dose estimates for an appreciable number of participants (the out-of-area group). Two sets of analyses were performed:

- The first set of analyses used single dose estimates for each living evaluable participant. In particular the median dose estimates (d_i as defined in section VIII.B.3.a above) were used as the primary point estimate of each participant's thyroid radiation dose from Hanford's ^{131}I. This approach, which is generally analogous to that used for the main published analysis of results from the Utah Study (134), ignored the uncertainty of the dose estimates and might therefore be expected to introduce bias into estimation of dose-response relationships. However this analysis had several advantages: it corresponded to the manner in which dose-response relationships were displayed in the tabular and graphical formats; it was analogous to analyses that have been performed for other studies of the effects of radiation on thyroid disease; and it was expected to provide reasonably accurate significance test results (as observed empirically [134]).

- The second set of analyses investigated the effects of uncertainty in the estimated doses from Hanford's ^{131}I. It was expected that this would have little impact on the results of significance tests of dose-response relationships.

These two sets of analyses are described in more detail in the following section.

C.2.a. Analyses Ignoring Dose Uncertainties

For the analyses ignoring dose uncertainties, the primary method for calculating parameter estimates was the method of maximum likelihood. In addition, however, certain analyses were performed using an alternative method, the method of least squares. Maximum likelihood and least squares are two generally applicable methods for estimating parameters of statistical models, including dose-response models. The specific implementations of these methods for HTDS are described in the following sections.

C.2.a.1. Maximum Likelihood Analyses of the Sex-stratified Linear Probability Model

Maximum likelihood estimates were calculated for the sex-stratified linear probability model [1] for disease outcomes and thyroid UDAs. As described in section VIII.C.1.a, the possible values of the sex-specific background rates A_1 and A_2 and of the slope B are constrained by the requirement that probabilities $P_j(d)$ must lie between 0 and 1. For example, if the background disease rates are low and there are few or no cases at the high end of the dose range, the regression parameter B in [1] is likely to be negative (e.g., $B < 0$). If it is too negative, there will be some study participants whose estimated probabilities are less than 0. To reduce difficulties that could arise from attempting to maximize the likelihood function under this constraint, estimation was based on the profile likelihood function. Let $L(A_1, A_2, B)$ denote the log-likelihood function, i.e., the logarithm of the probability of the observed outcome data given the doses and parameter values A_1, A_2, and B:

$$[7] \qquad L(A_1, A_2, B) = \Sigma_i \{ Y_i \times \ln[P_{j(i)}(d_i)] + (1 - Y_i) \times \ln[1 - P_{j(i)}(d_i)] \}$$

where d_i denotes the estimated thyroid radiation dose of participant i, j(i) is the stratum of participant i, and Y_i is an indicator variable for the outcome of interest, i.e., $Y_i = 1$ if the disease of thyroid UDA is detected in participant i, and $Y_i = 0$ otherwise. In [7] the summation is taken over $i = 1, \ldots, N$ where N is the number of living evaluable in-area participants. Note that the parameters to be estimated enter [7] through the sex-stratified linear model for $P_{j(i)}(d)$, i.e., equation [1]. For a given value of the slope parameter B, the

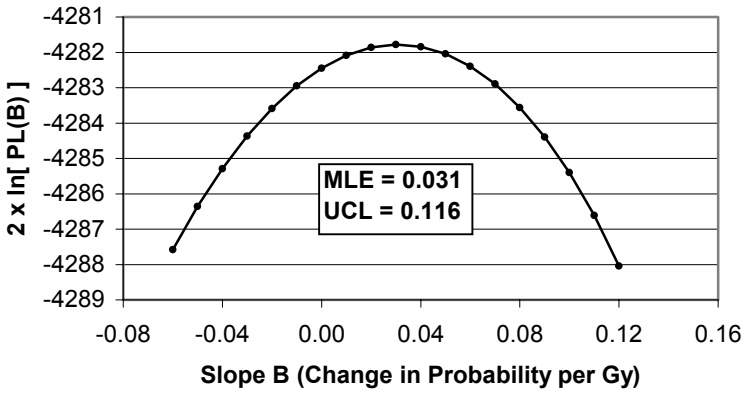
log-likelihood [7] varies as a function the two background rates A_1 and A_2. Let $A_{1,max}(B)$ and $A_{2,max}(B)$ denote the values of A_1 and A_2 for which $L(A_1, A_2, B)$ is maximized (again for the given value of B). Then the profile log-likelihood function can be written as

$$PL(B) = L[A_{1,max}(B), A_{2,max}(B), B] .$$

Note that the profile log-likelihood is simply a function of a single parameter, the slope B.

The maximum likelihood estimates of all three parameters of the sex-stratified linear probability model can be obtained by finding B_{MLE}, the value of B for which PL(B) is maximized. The maximum likelihood estimates of the background rates are then simply $A_{j,MLE} = A_{j,max}(B_{MLE})$ for j = 1 and 2. Figure VIII.C-1 below illustrates the profile log-likelihood function for the outcome of any thyroid ultrasound-detected abnormality (UDA; see section IX.P.2 below for the complete analysis of this outcome).

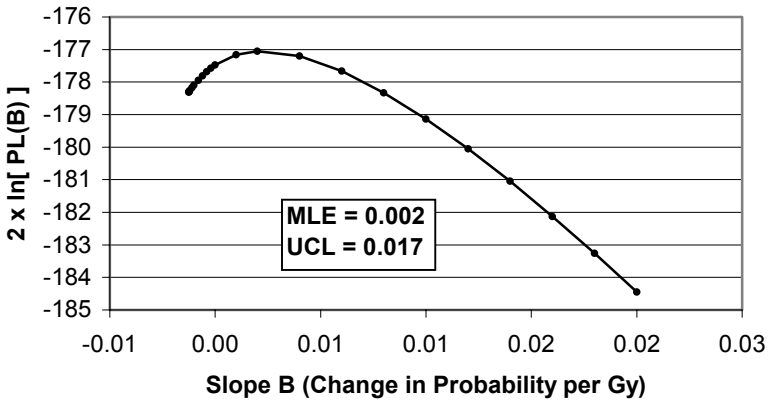
Figure VIII.C-1. Profile Log-likelihood Function for Any Thyroid UDA



The vertical axis displays the value of the natural logarithm of the profile likelihood function, PL(B), multiplied by 2, since this value is used for significance testing and confidence interval calculation as described below. MLE = maximum likelihood estimate, UCL = upper confidence limit (see section VIII.C.2.b.1 below).

Note that since conversion to logarithms and multiplication by 2 are monotone increasing transformations, the value of B which maximizes the curve in Figure VIII.C-1 also maximizes the profile likelihood function itself, and is therefore the maximum likelihood estimate of the slope. The example of any thyroid UDA in Figure VIII.C-1 above is comparatively well-behaved. That is, the maximum of the profile likelihood function is clearly evident, with values decreasing sharply and fairly symmetrically for either smaller or larger values of the slope B. This occurred because the background prevalence of any thyroid UDA was relatively high. Therefore the requirement that probabilities lie between 0 and 1 imposed no practical constraint. The situation was somewhat different for disease outcomes with low background probabilities. Figure VIII.C-2 displays the profile log-likelihood function for the outcome of thyroid cancer (see section IX.C below for the complete analysis of this outcome).

Figure VIII.C-2. Profile Log-likelihood Function for Thyroid Cancer

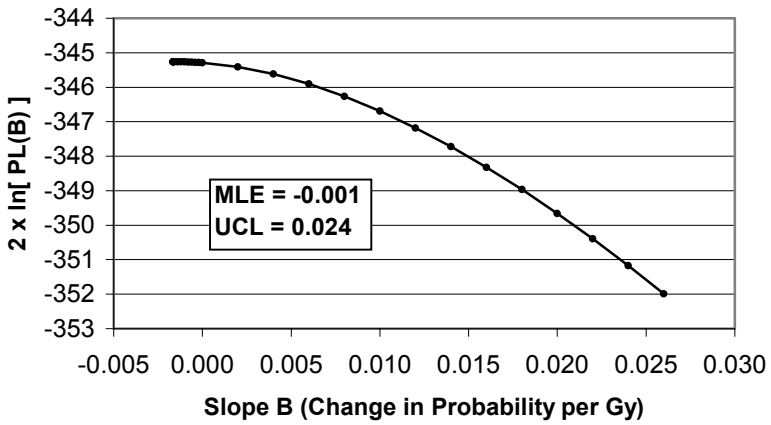


The vertical axis displays the value of the natural logarithm of the profile likelihood function, PL(B), multiplied by 2, since this value is used for significance testing and confidence interval calculation as described below. MLE = maximum likelihood estimate, UCL = upper confidence limit (see section VIII.C.2.b.1 below).

In contrast to the relatively high prevalence of thyroid UDAs, the background probabilities of thyroid cancer are low: only 14 (0.4%) of the 3191 living evaluable in-area participants had diagnoses of thyroid cancer based on the primary diagnostic definition (see section IX.C below). Therefore, due to the constraint described above, the profile likelihood function begins to decrease precipitously for increasingly negative values of the slope B. Nevertheless the value of B for which the profile likelihood is maximized is clearly evident.

Figure VIII.C-3 below displays the profile log-likelihood function for the outcome of Graves disease, which like thyroid cancer is relatively uncommon: 32 (1.0%) of the 3191 living evaluable in-area participants had diagnoses of Graves disease based on the primary diagnostic definition (see section IX.I. below). In this case the maximum likelihood estimate of B was slightly negative, and the profile likelihood function decreases very rapidly for even slightly more negative values of B.

Figure VIII.C-3. Profile Log-likelihood Function for Graves Disease



The vertical axis displays the value of the natural logarithm of the profile likelihood function, PL(B), multiplied by 2, since this value is used for significance testing and confidence interval calculation as described below. MLE = maximum likelihood estimate, UCL = upper confidence limit (see section VIII.C.2.b.1 below).

Computer programs were written by HTDS staff to fit models of the sex-stratified linear probability model [1] using the Newton-Raphson method to maximize the profile likelihood function. Likelihood ratio tests based on the profile likelihood function were used to test the statistical significance of the dose-response relationships in these analyses. Specifically, the test statistic was

$$\chi^2 = 2 \times \{ \ln[PL(0)]) - \ln[PL(B_{MLE})] \} ,$$

which, under the null hypothesis that the slope is 0, has a chi-square distribution with 1 degree of freedom. P-values for testing the one-sided alternative hypothesis that risk increases with increasing dose were calculated as

$$P = [1 - F_1(\chi^2)] / 2, \text{ if } B_{MLE} \geq 0$$

$$[1 + F_1(\chi^2)] / 2, \text{ if } B_{MLE} < 0 ,$$

where $F_1(\bullet)$ is the cumulative distribution function of the chi-square distribution with 1 degree of freedom. Confidence intervals for the parameters of the sex-stratified linear probability model were calculated as described in VIII.C.2.b.1 below.

C.2.a.2. Maximum Likelihood Analyses of Logistic Models

The parameters of logistic dose-response models of the form [2] and its generalizations for analyses of possible confounding and effect modification (see section VIII.C.1.b above) were estimated using SAS PROC LOGISTIC, a commercially available program for fitting logistic regression models.

Logistic regression models were also used to analyze disease outcomes in relation to alternative representations of exposure, i.e., by geostratum and by the dichotomous exposure variable described in section VIII.B.3.b above. For analyses by geostratum, each living evaluable participant was assigned a vector of eight indicator variables:

$G_{i,1}$ = 1 if participant i had geostratum Pasco/Kennewick, = 0 otherwise,

$G_{i,2}$ = 1 if participant i had geostratum Walla Walla City, = 0 otherwise, etc.

Note that these eight indicator variables all had value 0 for participants in the Richland geostratum. A similar indicator variable was defined for participants in the high or low exposure groups:

H_i = 1 if participant i is in the high exposure group, = 0 otherwise.

In both of these analyses age at the time of HTDS examination was included in the regression model. This was done because the nine geostrata differed slightly in their participants' average ages at examination, as did the high and low exposure groups (see section IX.A.7). Since cumulative incidence of thyroid diseases and prevalence of thyroid UDA increase with age, age at HTDS examination was included in order to adjust for its possible confounding effect. The sex-stratified age adjusted logistic regression models that allowed the cumulative incidence of disease outcomes or prevalence of thyroid UDA outcomes to vary among the geostrata or between the high and low exposure groups were then

$$P_{j(i)}(d_i) = \exp[\, A_{j(i)} + B_c \times C_i + \Sigma_g (B_g \times G_{i,g}) \,] \,/\, \{1 + \exp[\, A_{j(i)} + B_c \times C_i + \Sigma_g (B_g \times G_{i,g}) \,] \}$$

and

$$P_{j(i)}(d_i) = \exp(\, A_{j(i)} + B_c \times C_i + B \times H_i \,) \,/\, [1 + \exp(\, A_{j(i)} + B_c \times C_i + B \times H_i \,) \,] ,$$

respectively, where

C_i = age at HTDS examination for participant i.

C.2.a.3. Maximum Likelihood Analyses of Dose-Response Models for Laboratory Values, Thyroid Mass, and Numbers of UDAs

Since laboratory values and thyroid mass are quantitative variables, some of which are subject to censoring, their dose response models [4] were analyzed using SAS PROC LIFEREG, which calculates maximum likelihood estimates of parameters for parametric models. Maximum likelihood estimates were also calculated for the Poisson regression model used to analyze numbers of thyroid UDAs as described in section VIII.C.1.d above; the statistical program SPlus® was used for these calculations.

C.2.a.4. Least Squares Analyses

The method of least squares is another general method for estimating parameters in statistical models, an alternative to the method of maximum likelihood. SAS PROC REG was used to perform unweighted least squares analyses for three types of analysis of HTDS data. The first was to estimate the parameters of the sex-stratified linear probability model [1]. This analysis, designated the least squares analysis of ungrouped data ("LSU"), used the same data as in the profile likelihood analyses described in section VIII.C.2.a.1 above, i.e., the individual dose estimates and disease outcome data for each living evaluable participant.

For the second analysis, designated the least squares analysis of grouped data ("LSG"), the living evaluable in-area participants were grouped into eight dose categories, with cutpoints at 10, 50, 100, 150, 200, 300, and 400 mGy. The arithmetic means of the estimated thyroid doses of the participants in each category were then calculated separately for women and men. These average doses were then substituted for the individual dose estimates in the sex-stratified linear probability model [1].

The third use of least squares was to estimate the parameters of the linear-quadratic dose-response model [5].

C.2.b. Confidence Intervals

Confidence intervals were calculated for all estimates of background rates and dose-response slopes or regression coefficients. The method used to calculate confidence limits depended on the model and method of estimation, as described below. In addition, since three or more parameters were estimated for each model considered, confidence intervals were adjusted for the simultaneous estimation of multiple parameters as described in section VIII.C.2.b.4 below.

C.2.b.1. Linear Probability Model

As described above in section VIII.C.2.a.1, the parameters of the sex-stratified linear probability model [1] were estimated by maximizing the profile likelihood function. The profile likelihood function was also used to calculate confidence limits as follows:

$$B_{LCL} = \max\{\, B \mid B < B_{MLE} \text{ and } \ln[\, PL(B)\,] \le \ln[\, PL(B_{MLE})\,] - 0.5 \times Q(c_B\,,\, 1)\,\} \text{ and}$$
$$B_{UCL} = \min\{\, B \mid B > B_{MLE} \text{ and } \ln[\, PL(B)\,] \le \ln[\, PL(B_{MLE})\,] - 0.5 \times Q(c_B\,,\, 1)\,\}\,,$$

where

B_{LCL} is the lower confidence limit for the slope,
B_{UCL} is the upper confidence limit for the slope,
B_{MLE} is the maximum likelihood estimate of the slope,
c_B is the confidence level (see VIII.C.2.b.4 below), and
$Q(c_B\,,\, 1)$ is the c_B-th percentile of the chi-square distribution with 1 degree of freedom

Confidence limits for the two sex-specific intercepts were calculated as

[8] $A_{j,LCL} = A_{j,MLE} - Z[(100 + c_B)/2] \times SE(A_{j,MLE})$ and
$A_{j,UCL} = A_{j,MLE} + Z[(100 + c_B)/2] \times SE(A_{j,MLE})$, for j = 1, 2,

where

$A_{j,LCL}$ is the lower confidence limit for the sex-specific intercept A_j,
$A_{j,UCL}$ is the upper confidence limit for A_j,
$A_{j,MLE}$ is the maximum likelihood estimate of A_j,
$SE(A_{j,MLE})$ is the estimated standard error of $A_{j,MLE}$,
c_B is the confidence level (see VIII.C.2.b.4 below), and
$Z[(100 + c_B)/2]$ is the $[(100 + c_B)/2]$-th percentile of the standard normal distribution.

C.2.b.2. Logistic Models

Confidence intervals for parameters of logistic dose-response model [2] and its generalizations for analyses of confounding and effect modification were calculated from the parameters estimates and their estimated standard errors. For example, for the simple logistic dose-response model [2], confidence limits for the regression coefficient were calculated as

[9] $B_{LCL} = B_{MLE} - Z[(100 + c_B)/2] \times SE(B_{MLE})$ and
$B_{UCL} = B_{MLE} + Z[(100 + c_B)/2] \times SE(B_{MLE})$,

where

B_{LCL} is the lower confidence limit for the regression parameter,
B_{UCL} is the upper confidence limit for the regression parameter,
B_{MLE} is the maximum likelihood estimate of the regression parameter,
$SE(B_{MLE})$ is the estimated standard error of B_{MLE} ,
c_B is the confidence level (see VIII.C.2.b.4 below), and
$Z[(100 + c_B)/2]$ is the $[(100 + c_B)/2]$-th percentile of the standard normal distribution.

Confidence limits for the intercept parameters A_1 and A_2 in logistic models were calculated using [8], and converted into confidence limits for the background rates (see [3] above) as follows:

$P_{i,LCL}(0) = \exp(A_{j,LCL}) / [\ 1 + \exp(A_{j,LCL})\ ]$ and
$P_{i,UCL}(0) = \exp(A_{j,UCL}) / [\ 1 + \exp(A_{j,UCL})\ ]$ for j = 1, 2.

C.2.b.3. Models Fit by Method of Least Squares

For parameters estimated by the method of least squares, confidence intervals were calculated using [9] for the coefficients of dose terms and [8] for intercept terms. For linear-quadratic models, confidence intervals for the coefficients of the linear and quadratic terms were calculated as:

[9] $B_{t,LCL} = B_{t,MLE} - Z[(100 + c_B)/2] \times SE(B_{t,MLE})$ and
$B_{t,UCL} = B_{t,MLE} + Z[(100 + c_B)/2] \times SE(B_{t,MLE})$,

where

$B_{t,LCL}$ is the lower confidence limit for the coefficient of the linear (t=1) or quadratic (t=2) term,
$B_{t,UCL}$ is the upper confidence limit for the coefficient of the linear (t=1) or quadratic (t=2) term,
$B_{t,MLE}$ is the least sqaures estimate of the regression coefficient, for t = 1 or 2,
$SE(B_{t,MLE})$ is the estimated standard error of $B_{t,MLE}$,
c_B is the confidence level (see section VIII.C.2.b.4 below), and
$Z[(100 + c_B)/2]$ is the $[(100 + c_B)/2]$-th percentile of the standard normal distribution.

C.2.b.4. Confidence Level and Bonferroni Adjustment

The goal in calculating confidence intervals was to achieve a nominal 95% confidence level. However when confidence intervals with a given nominal confidence level are calculated simultaneously for more than one parameter of a model, the probability that all of the intervals contain the true values of their respective parameters is less than the nominal confidence level. For example, if 95% confidence intervals are to be calculated for each of the three parameters of the simple sex-stratified linear probability model [1] (i.e., the slope and the two sex-specific background rates), the probability that all three intervals will contain their true parameter values is less than 95%. In order to adjust for this effect of estimating multiple parameters, the Bonferroni method was used. In this method confidence intervals are calculated at a confidence level higher than the nominal level in order to ensure that the probability that all confidence intervals for a given model contain their respective true parameter values is not less than the nominal confidence level. Specifically, if confidence intervals are calculated for k parameters, then to achieve an overall confidence level no less than c, each confidence interval is calculated using confidence level

$c_B = 1 - (\ 1 - c\ ) / k$.

Thus for models with three parameters, in order to ensure overall confidence level no less than c = 95%, the three confidence intervals are each calculated at level $c_B = 98.33\%$. For models with four or five parameters, $c_B = 98.75\%$ or 99%, respectively.

Since $c_B > c$, each parameter's Bonferroni-adjusted confidence interval is wider than its unadjusted interval. In particular, upper confidence limits for slopes and dose-response regression parameters are higher with the Bonferroni adjustment than without.

C.2.c. Analyses of the Effect of Dose Uncertainties

It has long been recognized that the estimation of parameters in regression models such as the linear probability model [1] or the logistic model [2] is subject to bias if an independent variable (thyroid dose in this case) is observed with nonsystematic error, i.e., with an uncertainty that does not tend to systematically reduce or increase the values of the independent variable. In general, the effect of such error is to "attenuate" the estimate of the regression coefficient. That is, if the outcome variable tends to increase as the true value of the independent variable increases, the regression coefficient will tend to be underestimated. This phenomenon was observed by Kerber et al. (134) in their analysis of the Utah Thyroid Study. Significance tests can also be affected by error in the independent variable, although this is usually less of a problem.

A number of approaches have been devised over the years to deal with this problem, i.e., to correct or "deattentuate" estimates of regression coefficients. An approach analogous to that taken in the Utah Thyroid Study (133, 134) was used to calculate "deattentuated" estimates of the regression coefficient B.

C.2.c.1. *General Approach*

The general approach of the Utah Thyroid Study (133) was followed in reporting the results of dose-response analyses. That is, the main results were based on analyses that used point estimates of dose and ignored dose uncertainty. The additional analyses that adjusted for dose uncertainties were performed to illustrate how that adjustment affects the estimates and statistical significance of the dose-response relationships. Reporting results based on analyses that ignored uncertainties in the estimated thyroid doses from Hanford's ^{131}I was important for two reasons.

- The results of the conventional analyses (e.g., fitted dose-response functions based on the participants' median or mean doses) will be useful if one attempts to generalize from HTDS cohort to other persons with thyroid doses estimated by the CIDER program. The median dose estimate can then be applied to the corresponding HTDS estimates of dose-response functions from the conventional analysis. They cannot, however, be applied to the estimates based on the extended analysis that adjusts for uncertainty.

- The results of the conventional analysis will be comparable to the main results of the Utah Thyroid Study, which were reported in terms of mean doses (133).

C.2.c.2. Descriptive Analysis of Effects of Dose Uncertainty

To illustrate how the uncertainty of estimated doses influenced the fitted dose-response relationships, the linear dose-response models [1] were fit using each of the 100 realizations of dose separately. The point estimates of and Bonferroni-adjusted 95% confidence intervals for B, the slope parameter representing the magnitude of the radiation effect, were displayed graphically to illustrate how the estimate varied among the 100 realizations of dose, and how the estimates from the 100 realizations compare to those based on the average doses.

C.2.c.3. Estimation of B with Adjustment for Dose Uncertainty

A Bayesian approach was used to calculate "deattentuated" estimates of the regression coefficient B in the sex-stratified logistic model [2]. This approach specifies the relationship among observed data (outcomes and estimated thyroid doses), unobserved data (the participants' true doses), and the parameters of distributions governing the observed and unobserved data. Specifically, let GM_i and GSD_i denote the geometric mean and geometric standard deviation of the i-th participant's 100 dose estimates, as defined in section VIII.B.3. above, and let T_i denote the logarithm of the unobserved true dose ("true log dose") of participant i. The analysis was performed under the following assumptions.

(i) The logarithm of each participant's geometric mean dose, i.e., the arithmetic mean of the logarithms of his or her 100 dose estimates, is normally distributed with mean T_i and variance $\ln^2(GSD_i)$:

$$\ln(GM_i) \sim N(T_i , \ln^2(GSD_i)) \text{ for } i = 1, \ldots, N.$$

(ii) The true log doses T_i are themselves normally distributed with means and variances that differ between G subgroups of participants:

$$T_i \sim N(M_{g(i)} , V_{g(i)}) \text{ for } i = 1, \ldots, N,$$

where $g(i) \subset \{1, \ldots, G\}$ is the index of the subgroup containing participant i. The parameters of the underlying distributions of true log doses, i.e., M_g and V_g for $g = 1, \ldots, G$, are of course unknown and must be estimated.

(iii) Given the participant's true dose, the probability of the disease outcome or thyroid UDA outcome of interest is independent of his or her estimated dose, i.e., for each living evaluable in-area participant i= 1, …, N, the probability of the outcome is given by

$$\text{Prob}(Y_i \mid GM_i , T_i) = \exp(A_{j(i)} + B \times T_i) / [1 + \exp(A_{j(i)} + B \times T_i)] ,$$

where Y_i is the indicator of the outcome ($Y_i = 1$ if the participant has the outcome, otherwise $Y_i = 0$).

(iv) To complete the specifications necessary to implement the Bayesian approach, relatively uninformative prior distributions were assigned for the regression parameters (A_1, A_2, and B) and the means of the distributions of true log doses ($M_1, \ldots, M_G$), i.e., normal distributions with mean 0 and variance 10^6. Since the variance V of the distributions of true log doses is required to be greater than 0, its prior distribution was taken to be the gamma distribution with shape parameter 0.001 and scale parameter 10^6.

For assumption (ii), the subgroups of participants were defined so that they would be likely to have different underlying distributions of true log doses. In particular the subgroups were defined by geostrata and year of birth strata as follows:

Subgroup 1 (relatively high doses): Richland, Pasco/Kennewick, and Benton, Franklin, and Adams County geostrata, and 1940-1945 birth year strata.

Subgroup 2 (intermediate doses): Richland, Pasco/Kennewick, and Benton, Franklin, and Adams County geostrata, and 1946 birth year strata; or Walla Walla City or County geostrata, and 1940-1945 birth year strata.

Subgroup 3 (relatively low doses): Walla Walla City or County geostrata, and 1946 birth year strata; or Okanogan and Ferry/Stevens County geostrata (any birth year strata).

In addition, the distributions of true doses in (ii) were assumed to have common variance V, i.e., $V_g = V$ for g = 1, …, G.

The objective of the Bayesian approach was to estimate the posterior marginal distribution of the regression parameters, conditional on the observed data, i.e., on the values of Y_i, GM_i, GSD_i for i = 1, …, N. This was accomplished using the Gibbs sampling technique, as implemented by the freeware WinBUGS package (available at http://www.mrc-bsu.cam.ac.uk/bugs).

To begin the Gibbs sampling, initial values were specified as follows:

$B = 0$,
$A_j = \ln[\, b_j / (1 - b_j)\,]$ for j = 1 or 2, where b_j is the proportion of participants of sex j with the outcome, and
$T_i = \ln(GM_i)$ for i = 1, …, N .

The initial values of the means of the distributions of log true doses for the three subgroups were simply the means of $\ln(GM_i)$, which are summarized in the following table:

Table VIII.C-2 Description of Log True Doses for Subgroups 1– 3

Subgroup	No. of Living Evaluable In-Area Participants	Mean of $\ln(GM_i)$	Variance of $\ln(GM_i)$
1	2173	-2.27	3.33
2	646	-2.82	1.28
3	372	-4.41	2.51

Based on the variances of $\ln(GM_i)$ in the three subgroups, the initial value of the common variance V was defined as 2.5.

With these initial values, the Gibbs sampler was run for 2000 "burn-in" iterations, then for 5000 iterations to provide the estimated posterior marginal distribution of the regression parameters conditional on the observed data. In particular, the median of the 5000 values of B from its estimated marginal distribution was used to provide a "de-attenuated" estimate of the dose-response coefficient. In addition, the percentiles of that marginal distribution were used to provide an empirical confidence interval for the regression coefficient. Specifically, in order to obtain empirical confidence limits adjusted by the Bonferroni technique for the simultaneous estimation of three parameters (see section VIII.C.2.b.4 above), the percentiles at the 0.83% and 99.17% levels, i.e., the 41st and 4959th largest values of B, were defined as the confidence limits. Finally, a one-tailed empirical p-value was calculated as the proportion of the 5000 realizations for which the simulated value of B was less than 0.

C.2.c.4. Out-of-Area Participants

The approach described above applies only to analyses limited to the in-area participants, i.e., to those for whom the CIDER program provides 100 realizations of estimated dose. No attempt was made to calculate deattenuated estimates of the dose-response relationships using both the in-area and out-of-area participants.

D. Exposures from the Nevada Test Site

The ability to estimate thyroid doses caused by fallout ^{131}I from the Nevada Test Site (NTS) became available during the course of HTDS, as described in section VIII.B.2.h above. This section describes the way that information about exposures to fallout ^{131}I from the NTS was used in the HTDS.

D.1. General Approach

The general approach was to treat exposure to ^{131}I from the NTS as a potential confounding factor or effect modifier. Therefore the primary analyses of exposure-outcome relationships remained as described above. Moreover the analysis of potential confounding and effect modification by NTS exposures was performed basically as described for other potential confounders. However some special steps were necessary for analyses involving the NTS exposures; these steps are described below.

The decision to treat NTS exposure as a potential confounding or effect-modifying factor meant that certain other analyses that might be considered possible were not performed.

- No attempt was made to estimate, or test the statistical significance of, dose-response relationships between thyroid disease outcomes (or other response variables) and estimated NTS doses. This was because the HTDS cohort was defined to provide adequate statistical power for investigating the effects of Hanford doses, not NTS doses. Therefore it was very likely to be inadequate for the latter purpose.

- No analyses were conducted in which estimates of Hanford and NTS doses were added together or otherwise combined for use as the exposure variable. There were two reasons for this: 1) the objectives of HTDS refer specifically to the effects of doses from Hanford, and 2) it is not clear that dose estimates from the two dosimetry systems are comparable, and is it not known how to combine the estimates of uncertainty of the two doses.

D.2. Handling of Disease Outcome Variables in Analyses Involving NTS Doses

Exposures to ^{131}I from Hanford and the NTS occurred over a prolonged period of time. Therefore careful consideration was given to the handling of outcomes that were determined while exposure was still occurring.

No special handling was necessary to accommodate NTS exposures in the analyses of the prevalence of thyroid UDAs, since these were based on the HTDS examination, long after the cessation of Hanford and NTS exposures. However, for disease outcomes the situation was different, since diagnoses might have occurred before the end of 1957, i.e., before the end of the period for which estimated doses from NTS fallout were calculated (see section VIII.B.2.h above). As described in section VIII.C.1.a above, the primary analysis of each disease outcome was based on the cumulative incidence of the outcome among living evaluable participants. "Cumulative incidence" referred specifically to the percentage of living evaluable participants with the outcome of interest diagnosed at any time up to and including the HTDS examination.

One possible alternative to the use of cumulative incidence was to perform an "incidence study," i.e., to analyze incidence rates (cases per 100,000 person-years). The use of incidence rates is a standard epidemiological method that has been successfully applied in a number of studies of radiation effects. It is often particularly appropriate when exposure occurs over a prolonged period. However when the HTDS protocol and Analysis Plan were developed (i.e., prior to the availability of information about exposures to ^{131}I from the NTS), two major reasons were identified that argued against trying to perform an incidence study.

- Dates of many past diagnoses (i.e., diagnoses made prior to the participant's HTDS examination) were likely to be imprecisely known. This was expected to be true to varying degrees for diagnoses documented in medical records. Reports of past diagnoses based solely on the In-Person Interview of the participant or the CATI were likely to be especially imprecise.

- The age profiles of incidence rates were likely to be distorted by the occurrence of HTDS examinations. NCRP has noted that tumor registries may underestimate the true incidence of thyroid cancer by a factor of three (36). Thus it was expected that the highly sensitive HTDS examinations would induce an apparent sharp jump in age-specific incidence rates of neoplastic diseases. This was anticipated for non-neoplastic diseases as well.

  In addition, a third enabling reason was identified. While this reason did not argue against analyzing incidence rates, it did provide a rationale for the use of cumulative incidence.

- The bulk of the Hanford exposure occurred before 1950, while the vast majority of diagnoses likely occurred later. Therefore an analysis of cumulative incidence, which in effect treats all diagnoses as occurring after completion of exposure, was unlikely to be seriously biased.

When the decision was made to include the NTS exposure as a potential confounding or effect modifying factor, the applicability of these reasons was re-examined. The first two reasons remained strong arguments against conducting an incidence study. However the third reason was more problematic: since the NTS exposures occurred primarily between 1952 and 1957, the likelihood that some diagnoses occurred before the end of exposure was increased (HTDS participants were 11 to 18 years old at the end of 1957). Therefore the decision was made to modify analyses that include NTS exposure by considering cumulative incidence since January 1, 1958. This was accomplished for each disease outcome by excluding any participant with a diagnosis of that outcome before January 1, 1958. Implications of this decision included the following:

- Restricting the period of observation to begin on January 1, 1958, rather than at birth, might have an impact on the estimates and/or statistical significance of the dose-response relationships of interest, i.e., the associations between cumulative incidence of disease outcomes and Hanford dose. It was considered likely that any such impact would be small, since it is expected that few diagnoses occurred before 1958. For any disease outcome with no diagnoses before 1958, the modification would have no effect. For disease outcomes that included diagnoses before 1958, the effect of restricting the period of observation was examined. The age-stratified linear probability model (equation [1] above) was fit using both the unrestricted and the restricted periods of observation. In addition the generalized linear probability model including the pooled categorical variable for NTS exposure was also fit using both the unrestricted and the restricted periods of observation. The results of these four fits were compared to assess whether the choice of observation period affects the estimated effect of Hanford doses.

- To perform the modified analysis, each diagnosis had to be classified according to its date: before 1958 or after 1957. As noted in Reason 1 above, dates of many past diagnoses were known only imprecisely. However many cases with imprecisely known diagnosis dates could be accurately assigned between these two time intervals. Only a few cases could not be accurately assigned with a high degree of certainty. Each such case was assigned a diagnosis date in the middle of the range of plausible dates based on the available information, and then into one of the two time intervals on the basis of that assigned date.

IX. RESULTS

A. Characteristics of the Living Evaluable Participants

Of the 3447 eligible participants who attended an HTDS clinic, seven (0.2%) were determined to be nonevaluable according to the criteria in section IV.B above. Six did not have complete residence histories for the period from the beginning of their possible exposure to ^{131}I from Hanford through the end of 1957, and the seventh had a tracheotomy tube in place which prevented palpation of her thyroid at her HTDS clinical examination. The remaining 3440, designated the living evaluable participants, are the basis for most of the analyses reported here. For each of these 3440 participants, sufficient data were available to permit an evaluation of thyroid health and estimation of the radiation dose from Hanford's ^{131}I. Several characteristics of the living evaluable participants are summarized in Table IX.A-1. About half (50.8%) of the living evaluable participants were women. About one-fourth (26.3%) were born in 1944, and another third (34.0%) were born in 1943 or 1945. Therefore a large proportion of the living evaluable participants were infants or very young children during 1945, the years of the largest releases of ^{131}I from Hanford. At the time of their HTDS examinations, the living evaluable participants ranged in age from 45 to 57 (median 51). A large majority (97.5%) described themselves as white or Caucasian.

Table IX.A-1. Characteristics of Living Evaluable Participants

Characteristic		No.	%
Sex	Female	1747	50.8
	Male	1693	49.2
	Total	3440	100.0
Year of birth	1940	243	7.1
	1941	283	8.2
	1942	472	13.7
	1943	560	16.3
	1944	906	26.3
	1945	611	17.8
	1946	365	10.6
	Total	3440	100.0
Age at examination	45	1	--
	46	58	1.7
	47	194	5.6
	48	264	7.7
	49	323	9.4
	50	388	11.3
	51	741	21.5
	52	561	16.3
	53	278	8.1
	54	229	6.7
	55	273	7.9
	56	118	3.4
	57	12	0.3
	Total	3440	100.0
Race/ethnic origin	White/Caucasian	3354	97.5
	Black/Negro	1	--
	Asian or Pacific Islander	10	0.3
	Native American	40	1.2
	Spanish or Hispanic	5	0.1
	Other	23	0.7
	Don't Know	2	0.1
	Not Recorded	1	--
	Refused	4	0.1
	Total	3440	100.0
Religious preference	Protestant	2176	63.3
	Catholic	483	14.0
	Jewish	4	0.1
	Mormon	128	3.7
	Seventh Day Adventist	108	3.1
	Other	94	2.7
	None	437	12.7
	Not Recorded	1	--
	Refused	6	0.2
	Don't Know	3	0.1
	Total	3440	100.0

One important purpose for collecting information about the characteristics described above (as well as the other factors described below) was to use that information to test for possible confounding and effect modification in the analyses of the radiation dose-responses (see section VIII.A.1.b above). As noted above, the living evaluable participants overwhelmingly identified themselves as white or Caucasian (97.5%). Therefore, meaningful analyses of race or ethnic origin as a potential confounder or effect modifier could not be performed. In addition, Jewish religious preference was of particular interest as a potential confounder or effect modifier, since there is some evidence suggesting increased risk of thyroid cancer in Jewish populations (136). However only four (0.1%) of the living evaluable participants stated Jewish as their religious preference, so further analysis of this factor was not possible.

A.1. History of Diagnostic X-Rays, Fluoroscopy, Thyroid Nuclear Scans, and other Nuclear Medicine Procedures

Of the 3440 living evaluable participants, 3318 (96.4%) had a report, either from the In-Person Interview or the CATI, of one or more diagnostic x-rays or fluoroscopies of the upper body, thyroid nuclear scans, or other nuclear medicine procedures. The proportions with reports of specific procedures are summarized in Table IX.A-2.

Table IX.A-2. History of Diagnostic X-Rays, Fluoroscopy, Thyroid Nuclear Scans, and other Nuclear Medicine Procedures

	In-Person Interview Only		In-Person Interview and CATI	
Have You Ever Had:	No.	%♦	No.	%♦
CAT scan of the upper body?	832	24.2	*	
Diagnostic x-rays of the head?	1183	34.4	1294	37.6
Diagnostic x-rays of the neck?	1026	29.8	1045	30.4
Diagnostic x-rays of the chest or upper body, including mammograms?	3027	88.0	3045	88.5
Diagnostic x-rays of the stomach or mid-back?	745	21.7	*	
Barium enema?	887	25.8	*	
Upper GI?	1228	35.7	1236	35.9
Intravenous pyelogram or IVP?	420	12.2	425	12.4
Fluoroscopy of the upper body?	234	6.8	271	7.9
Other nuclear scan?	229	6.7	231	6.7
Any of the above?	3305	96.1	3317	96.4

* Question not asked in CATI

♦ Percent calculated in relation to number of living evaluable participants

By far the most common types of procedures were diagnostic x-rays of the chest or upper body, including mammograms, which were reported for 3045 (88.5%) of the living evaluable participants. Also particularly common were diagnostic x-rays of the head or neck, which were reported for 37.6% and 30.4%, respectively, and upper GI examinations (35.9%). Nearly one-fourth (24.2%) of the living evaluable participants reported a history of upper body CAT scan (since the CATI covered the time period ending in 1957, it did not include a question regarding CAT scans). Upper body fluoroscopies were reported for 271 (7.9%) living evaluable participants. Histories of nuclear scans other than thyroid scans were reported for 231 (6.7%) of the living evaluable participants. In addition, histories of thyroid nuclear scans, which are used to assist in the diagnosis of thyroid disorders, were reported for 142 (4.1%) of the living evaluable participants.

A.2. History of Radiation Treatment

Histories of x-ray treatment affecting the upper body for reasons other than cancer were reported for 90 (2.6%) of the living evaluable participants (Table IX.A-3). The most common reason stated was treatment of acne, reported for 37 (1.1%).

Cancer other than thyroid cancer was reported on the In-Person Interview by 276 (8.0%) of the living evaluable participants, with 42 of these 276 reporting having received radiation therapy for the cancer.

Table IX.A-3. History of Radiation Treatment

	In-Person Interview Only		In-Person Interview and CATI	
Have You Ever Had:	No.	%♦	No.	%♦
X-ray treatment to the upper body for acne?	32	0.9	37	1.1
X-ray treatment to the upper body for ringworm?	1	0.03	10	0.3
X-ray treatment for enlarged tonsils?	2	0.06	4	0.1
X-ray treatment to the upper body for tuberculosis?	2	0.06	2	0.06
X-ray treatment for scalp infection?	1	0.03	1	0.03
X-ray treatment for enlarged thymus?	0	--	7	0.2
X-ray treatment to the upper body for any other reason?	15	0.4	31	0.9
Any of the above x-ray treatments?	52	1.5	90	2.6
History of any cancer other than thyroid?	276	8.0	*	
Radiation treatment for any cancer other than thyroid?	42	1.2		

* Question not asked in CATI

♦ Percent calculated in relation to number of living evaluable participants

A.3. History of Dental X-rays

A history of dental x-ray exposure was reported for nearly all (99.2%) of the living evaluable participants, although only 346 (10.1%) reported receiving dental x-rays more frequently than once per year during their life (see Table IX.A-4). About half (51.8%) of the living evaluable participants reported on the In-Person Interview or CATI at least one time period when lead shielding of the neck was not used in dental x-ray examinations.

Table IX.A-4. History of Dental X-rays

	In-Person Interview Only		In-Person Interview and CATI	
Have You Ever Had:	No.	%♦	No.	%♦
Dental x-ray?	3406	99.0	3414	99.2
Dental x-ray more frequently than once a year?	324	9.4	346	10.1
Dental x-rays that did not usually include a lead shield over the neck area?	1727	50.2	1781	51.8

♦ Percent calculated in relation to number of living evaluable participants

A.4. Occupational History

Data regarding the participants' occupational histories were obtained from the In-Person Interview. The intention was to identify persons who had worked in occupations that might involve exposure to radiation. Therefore, results are presented for occupations in the metals industry, employment at nuclear facilities, and other occupations that might involve exposure to radioactive materials. The results are summarized in Table IX.A-5.

Table IX.A-5. Occupational History

Have You Ever Worked in Any of the Following Industries or Occupations?	Female		Male		Total	
	No.	%♦	No.	%♦	No.	%♦
Any Metal Industry	30	1.7	239	14.1	269	7.8
Geology	5	0.3	19	1.1	24	0.7
Metallurgy	7	0.4	45	2.7	52	1.5
Metal processing	15	0.9	131	7.7	146	4.2
Ore refining	2	0.1	29	1.7	31	0.9
Mining	1	0.1	60	3.5	61	1.8
Any Nuclear Facility	110	6.3	273	16.1	383	11.1
Nuclear Industry, as a civilian	84	4.8	168	9.9	252	7.3
On the premises of a nuclear facility	85	4.9	224	13.2	309	9.0
Any Area Exposed to Radioactive Materials/X-Rays	203	11.6	274	16.2	477	13.9
Health care	178	10.2	81	4.8	259	7.5
Scientist, researcher, or student	34	1.9	56	3.3	90	2.6
Military	2	0.1	123	7.3	125	3.6
Any other industry or occupation	16	0.9	66	3.9	82	2.4
Any of the Above Industries	321	18.4	636	37.6	957	27.8

♦ Percent calculated in relation to number of living evaluable participants

Of the 3440 living evaluable participants, 957 (27.8%) reported a history of employment in one or more of the occupations or facilities of interest. The proportion was higher among the men (37.6%) compared to the women (18.4%). The higher proportion among the men was largely due to occupations in the metals industry (14.1%) or employment at nuclear facilities (16.1%). The proportions of women reporting such histories were much smaller (1.7% and 6.3% respectively). The proportions of men and women reporting histories of working in areas with possible exposure to radioactive materials were more similar: 16.2% for men, 11.6% for women. Most of the women with such histories identified an occupation in the health care industry (10.2%).

A.5. Smoking History

The In-Person Interview included a series of questions regarding smoking. Of the 3440 living evaluable participants, 2053 (59.7%) reported a history of ever smoking cigarettes, cigars, and/or pipes. See Table IX.A-6.

Table IX.A-6. Smoking: History of Ever Smoking

Have You Ever Smoked Any of the Following:	Female No.	Female %♦	Male No.	Male %♦	Total No.	Total %♦
Non-filter cigarettes?	180	10.3	755	44.6	935	27.2
Filter cigarettes?	880	50.4	1013	59.8	1893	55.0
Any cigarettes?	897	51.3	1103	65.2	2000	58.1
Cigars?	4	0.2	121	7.1	125	3.6
Pipe?	2	0.1	221	13.0	223	6.5
Any of the above?	897	51.3	1156	68.3	2053	59.7

♦ Percent calculated in relation to number of living evaluable participants

The proportion that reported ever smoking was higher among men (68.3%) than women (51.3%). As expected, cigarette smoking was by far the most common form, reported by all of the women who reported any kind of smoking, and by 65.2% of the 1693 living evaluable male participants. The amount of cigarette smoking was quantified in terms of pack-years. One pack-year is equivalent to smoking one pack a day for one year. When adequate data were available from the In-Person Interview, the total consumption in pack-years was calculated by integrating the reported consumption in packs per day over the years of smoking. Similar calculations were performed to quantify cigar and pipe smoking based on cigars per day and pipes (bowls) smoked per day. The results are shown in Table IX.A-7.

Table IX.A-7. Smoking: Level of Use

		Female	Male	Total
Non-filter cigarettes (pack-years)	Median	1.5	4	3
	Range	0.01-70	.0003-98.9	.0003-98.9
	Number	180	745	925
Filter cigarettes (pack-years)	Median	18.54	19	18.75
	Range	0.01-140	0.01-130	0.01-140
	Number	874	1009	1883
Any cigarettes (pack-years)	Median	19.5	24.1	21.25
	Range	0.01-140	0.01-136	0.01-140
	Number	891	1090	1981
Cigars (cigar-years)	Median	1	6	6
	Range	0.8-2	0.01-443	0.01-443
	Number	4	119	123
Pipe (pipe-years)	Median	1.1	5.7	5.7
	Range	0.7-1.4	0.1-500	0.1-500
	Number	2	217	219

Among 1981 living evaluable participants who ever smoked cigarettes and for whom adequate consumption data were available, the median total pack-years was 21.25 (range 0.01 to 140). The consumption was higher among men who smoked cigarettes, with a median of 24.1 pack-years, compared to 19.5 pack-years for women. Cigar- and pipe-smoking men reported median consumption levels of 6 cigar-years and 5.7 pipe-years. Very few women reported ever smoking cigars or pipes.

A.6. Dietary Factors

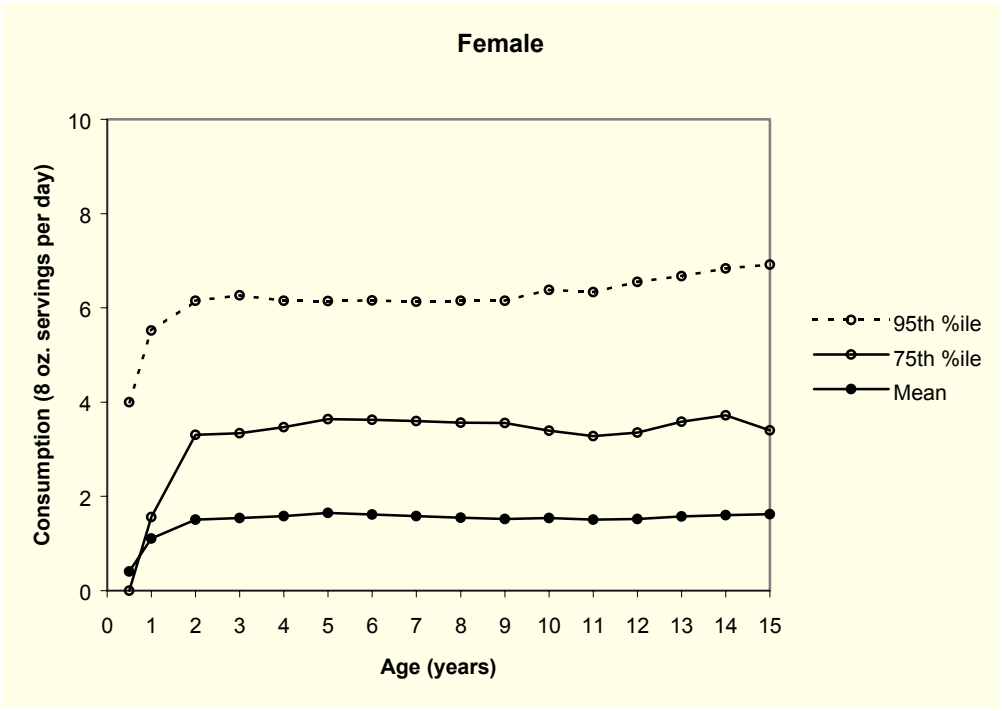
Each participant's thyroid radiation dose depends on several factors. Of particular importance is the dietary history, including the amounts of milk, milk products and other foods that the participant consumed, and the sources of those foods. Therefore, information about these dietary factors was collected as part of the CATI for use by the CIDER program in calculating estimated radiation doses. The following sections present information about the quantities of various milk and food products consumed by the 1979 living evaluable in-area participants whose CATI data were used for dose estimation, as reported by their CATI respondents. Results are shown separately for women and men. Since a participant's consumption of milk and food products typically changed over time, results are also shown by age. Specifically, descriptive statistics (mean, median, and 5^{th}, 25^{th}, 75^{th} and 95^{th} percentiles) were calculated for the reported amounts consumed on the participant's 6-month birthday and annual birthdays (first, second, etc.) through age 15. (Since CATI respondents often reported that dietary factors changed on birthdays, the values used for these analyses were in fact those reported for 5 days after the participant's birthday.) If a participant was out of area on one of these occasions, or his or her consumption was reported as unknown, then he or she was excluded from the distribution for that birthday. Similarly, if a participant's birthday occurred before December 1944 or after 1957, then he or she was excluded from the distribution for that birthday. If a participant was reported not to have consumed a given milk or food product on a given birthday, then the consumption was taken to be zero.

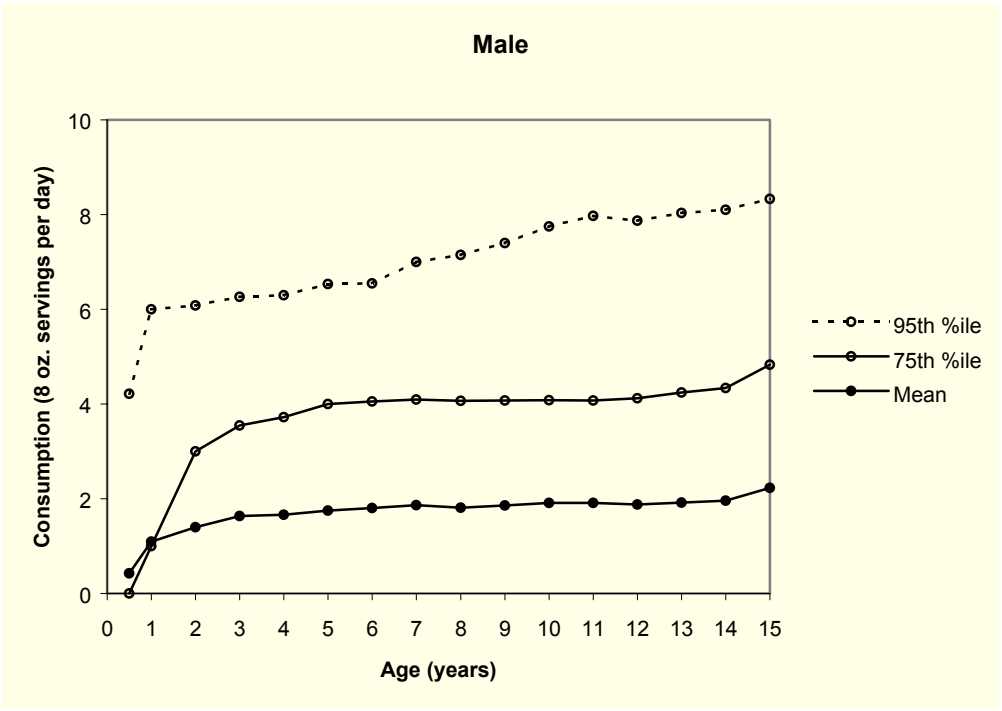
A.6.a. Raw Cow's Milk and Milk Products

Consumption of raw cow's milk or milk products was reported for 999 (50.5%) of the 1979 living evaluable in-area participants whose CATI data were used for dose estimation (498 women, 501 men). For 61 of these 975 participants (31 women, 30 men), the CATI respondent was unable to provide estimates of the quantity consumed. Figure IX.A-1 summarizes the distributions of raw cow's milk and milk products consumption by sex and age. Each participant's consumption, expressed as 8 oz. servings per day, was calculated from that reported for glasses of milk, other servings of milk, and milk products. At every age shown in the figures, fewer than half of the participants were reported to consume raw cow's milk or milk products, and the 5^{th} and 25^{th} percentiles and medians were consequently all zero; therefore these three statistics are omitted from the figure for clarity.

As shown in Figure IX.A-1, consumption of raw cow's milk and milk products increased sharply for both sexes until age 2, then leveled off at averages of about 1.6 and 1.9 eight oz.servings per day for women and men, respectively. For both sexes, only about 12% were reported to consume raw cow's milk or milk products at 6 months of age, and only about 28% at one year of age. At older ages the proportions of nonconsumers ranged between 33% and 40%.

Figure IX.A-1. Raw Cow's Milk Consumption, by Sex and Age



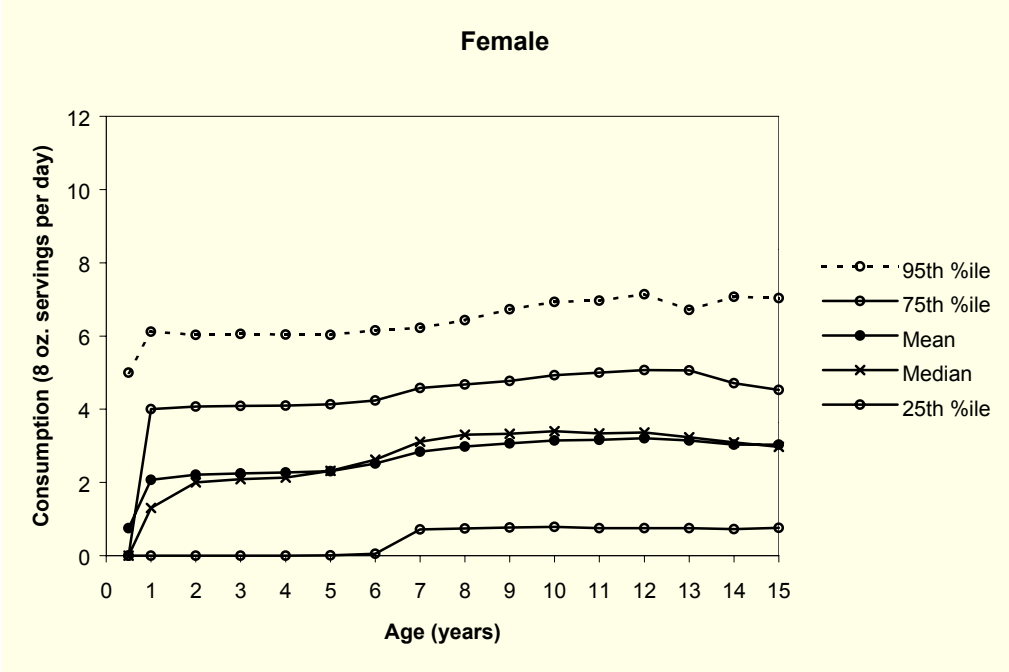


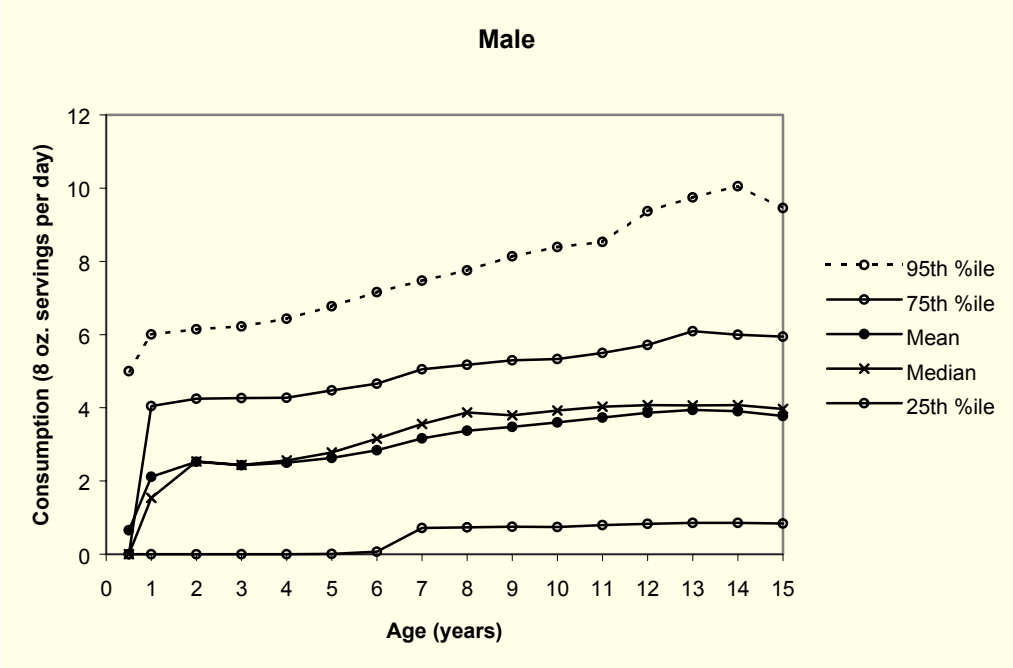
A.6.b. Processed Cow's Milk and Milk Products

Consumption of processed cow's milk or milk products was reported for 1741 (88.0%) of the 1979 living evaluable in-area participants whose CATI data were used for dose estimation (871 women, 870 men). For 86 of these 1746 (44 women, 42 men), the CATI respondent was unable to provide estimates of the quantities consumed. Figure IX.A-2 summarizes the distributions of processed cow's milk and milk products consumption by sex and age. Each participant's consumption, expressed as 8 oz. servings per day, was calculated from that reported for glasses of milk, other servings of milk, and milk products. At every age shown in the figures, more than 10% of the participants were reported to be nonconsumers of processed cow's milk or milk products, and the 5th percentiles were consequently all zero; therefore the 5th percentiles are omitted from the figure for clarity.

As shown in Figure IX.A-2, consumption of processed cow's milk and milk products increased to about 2 eight oz. servings per day at one year of age. For women, consumption remained fairly stable at this level until about age 5 or 6, then increased to about three 8 oz. servings per day. For men, consumption increased fairly steadily until the teenage years, to just under four 8 oz. servings per day. For both sexes, the proportion of nonconsumers decreased from nearly 40% at age 1 to 20% at age 6; thereafter the proportions remained fairly stable at 10 – 14%.

Figure IX.A-2. Processed Cow's Milk Consumption, by Sex and Age





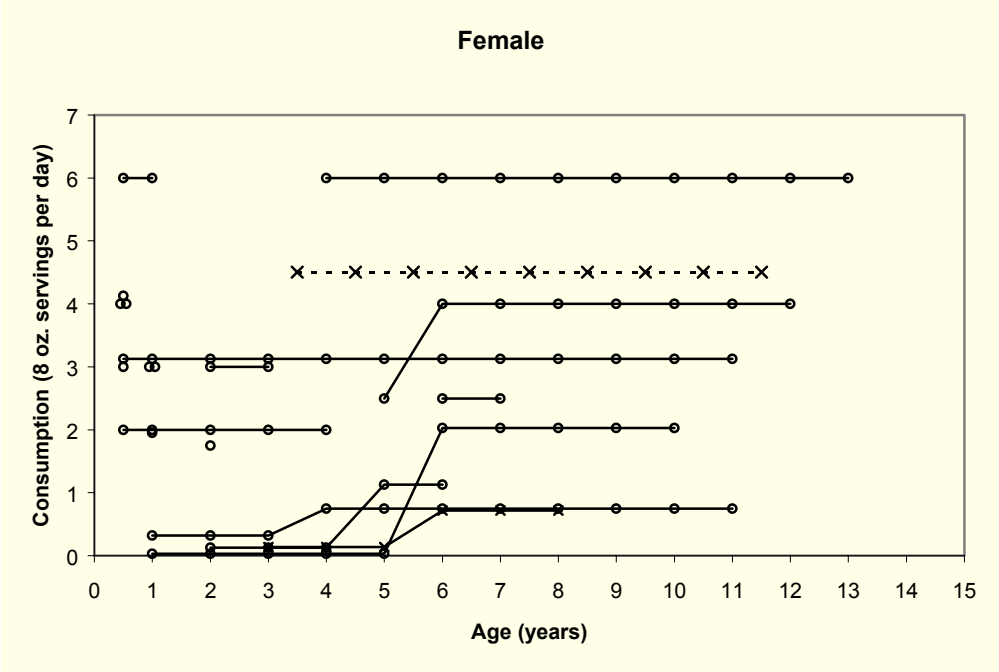
A.6.c. Goat's Milk and Milk Products

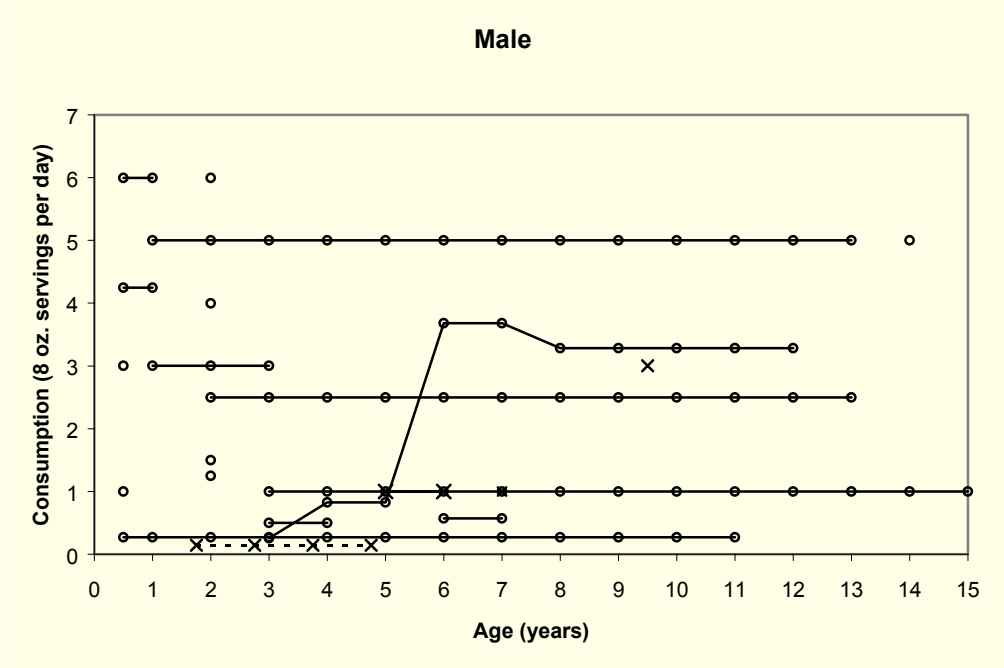
Consumption of goat's milk or milk products was reported for only 46 (2.3%) of the 1979 living evaluable in area participants whose CATI data were used for dose estimation (23 women, 23 men). For five of these 46 (three women, two men), the CATI respondent was unable to provide estimates of the quantities consumed. The reported consumption levels for the remaining 41 participants, i.e., the 20 women and 21 men for whom quantities consumed were reported, are displayed in Figure IX.A-3. Note that this figure differs from figures IX.A-1 and IX.A-2 above, since it displays the levels reported for individual participants rather than percentiles of the consumption levels.

In figure IX.A-3 below, the consumption levels reported for an individual participant at various ages are connected by lines. This was done to illustrate changes of consumption levels over time, and does not imply that every participant in the figure consumed goat's milk or milk products continuously throughout the ages indicated. For ages at which a participant was reported to be a nonconsumer of goat's milk or milk products, points are omitted from the figure for clarity. There were three participants for whom consumption of goat's milk or milk products was reported only for one or more periods of time that did not include their six-month birthday or any of their first through 15th birthdays. In order to include the participants in the figure, their consumption levels were plotted for ages at which consumption occurred. The reported consumption levels for these three participants are indicated in the figure by dashed lines and and "x" symbols. For example, one female participant was reported to have consumed about 4.5 servings of goat's milk or milk products per day between the ages of about 3.5 and 11.7 years, except during a two-month period of each year, a period that happened to include her birthday. Therefore her consumption levels are shown in Figure IX.A-3 below for ages 3.5 years, 4.5 years, etc.

The levels of goat's milk and milk product consumption reported by CATI respondents for these 46 participants ranged up to six 8 oz. servings per day for both women and men.

Figure IX.A-3. Goat's Milk Consumption, by Sex and Age



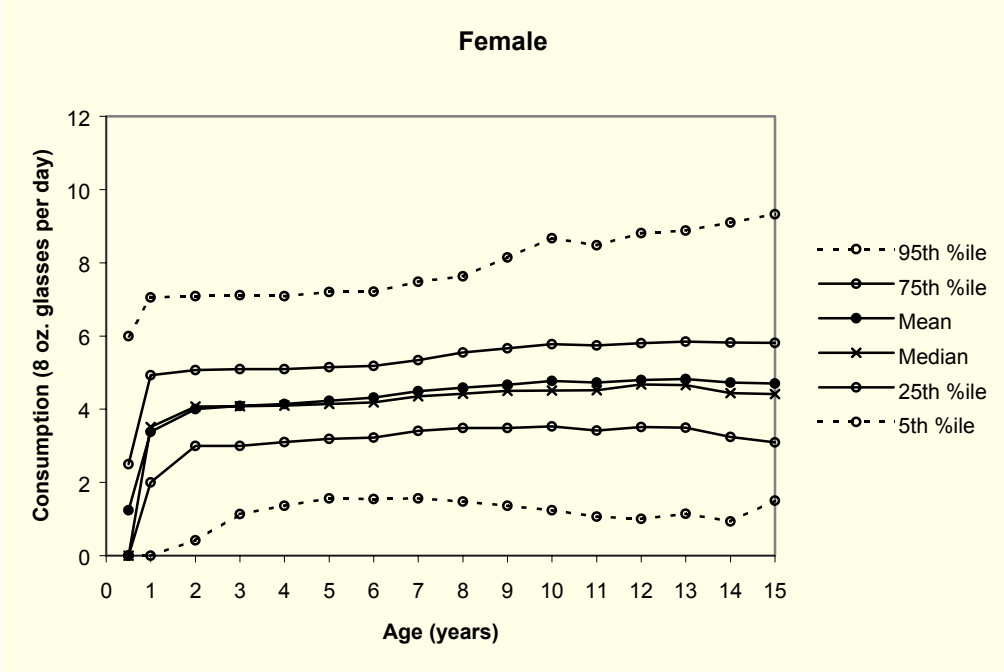


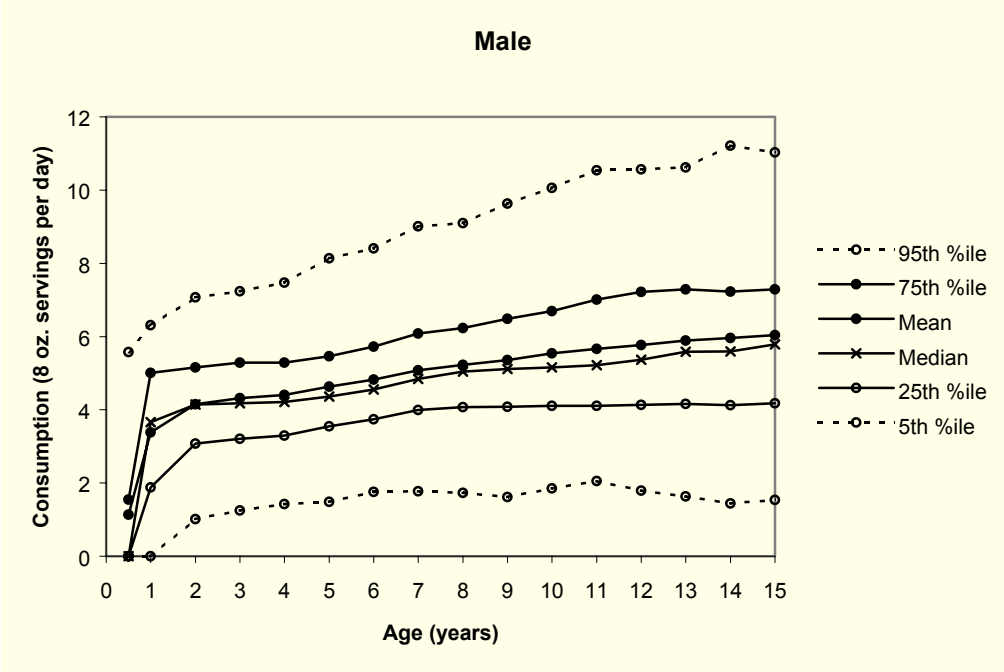
A.6.d. Total Milk and Milk Products

Figure IX.A-4 summarizes the distributions of total milk consumption (raw or processed cow's or goat's milk and milk products) by sex and age. Each participant's consumption, expressed as 8 oz. servings per day, was calculated from that reported for glasses of milk, other servings of milk, and milk products.

As shown in Figure IX.A-4, the reported consumption of milk and milk products increased to about four 8 oz. servings per day at two years of age. For women, consumption remained fairly stable, increasing slightly and gradually until the teenage years. For men, consumption increased steadily up to about six 8 oz. servings per day by age 15. For both sexes, the proportion for whom no milk consumption was reported was 69% at age 6 months and 15% at one year of age. For women, the proportion of nonconsumers fell to 3% at age 2 years, and was 2% or less for all older ages. For men, the proportion of nonconsumers was about 1% at 2 and 3 years of age, and 0.3% or less for all older ages.

Figure IX.A-4. Total Milk Consumption, by Sex and Age



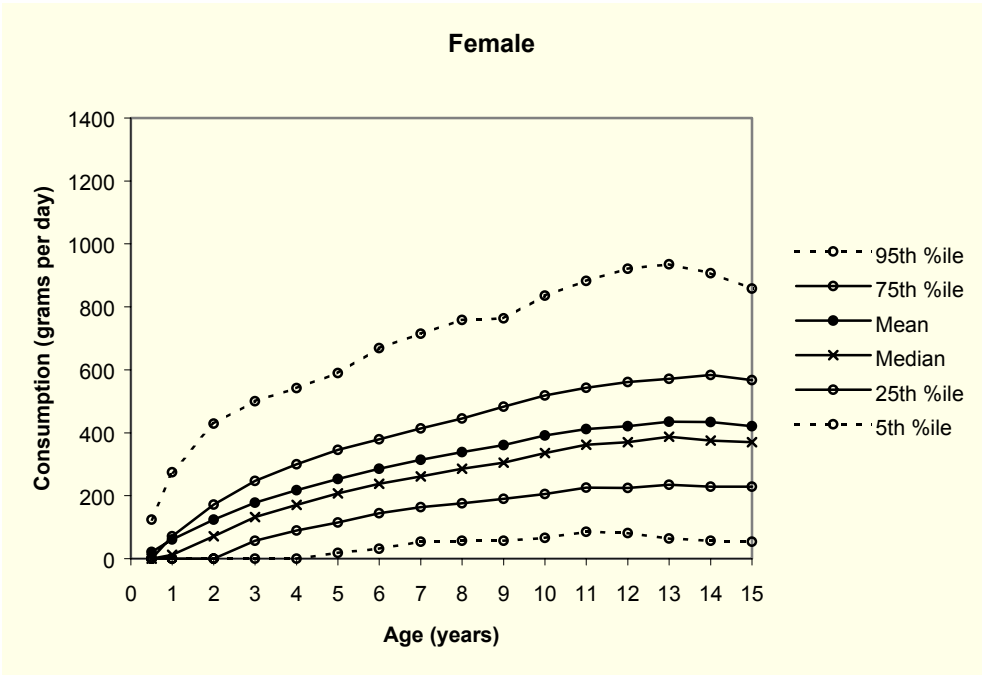


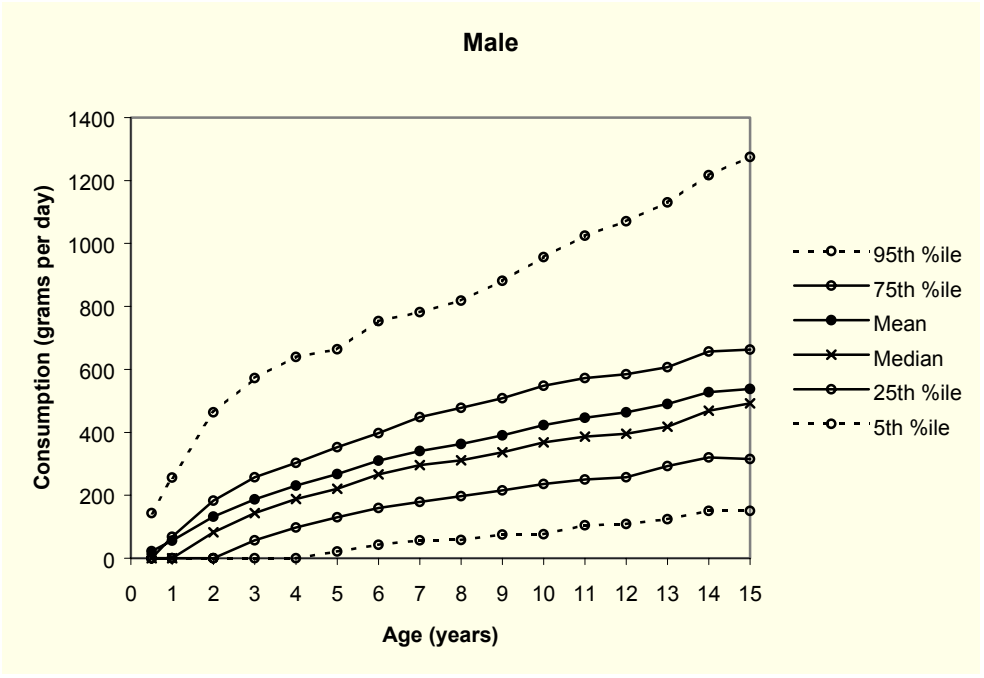
A.6.e. Fruit

Consumption of fruit was reported for 1786 (90.2%) of the 1979 living evaluable in-area participants whose CATI data were used for dose estimation (895 women, 891 men). For 144 of these 1786 (67 women, 77 men), the CATI respondent was unable to provide estimates of the quantities consumed. Figure IX.A-5 summarizes the distributions of fruit consumption by sex and age. Each participant's consumption, expressed as grams per day, was calculated from the information reported by the CATI respondent.

As shown in Figure IX.A-5, the reported consumption of fruit increased steadily with age for both sexes, to about 430 grams per day for women and over 535 grams per day for men by the teenage years. For both sexes, the proportion who were reported not to consume fruit decreased from about 75% at 6 months of age to about 50% and 25% at one and two years of age, respectively. The proportions continued to decrease with increasing age, reaching plateaus of about 2% for women and 1% or less for men after age 7.

Figure IX.A-5. Fruit Consumption, by Sex and Age



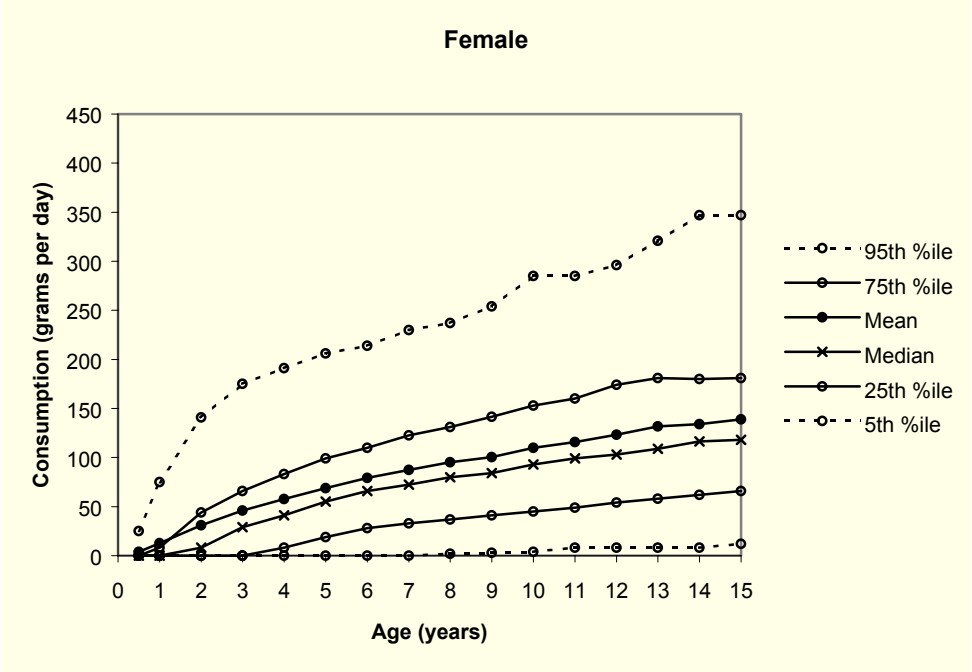


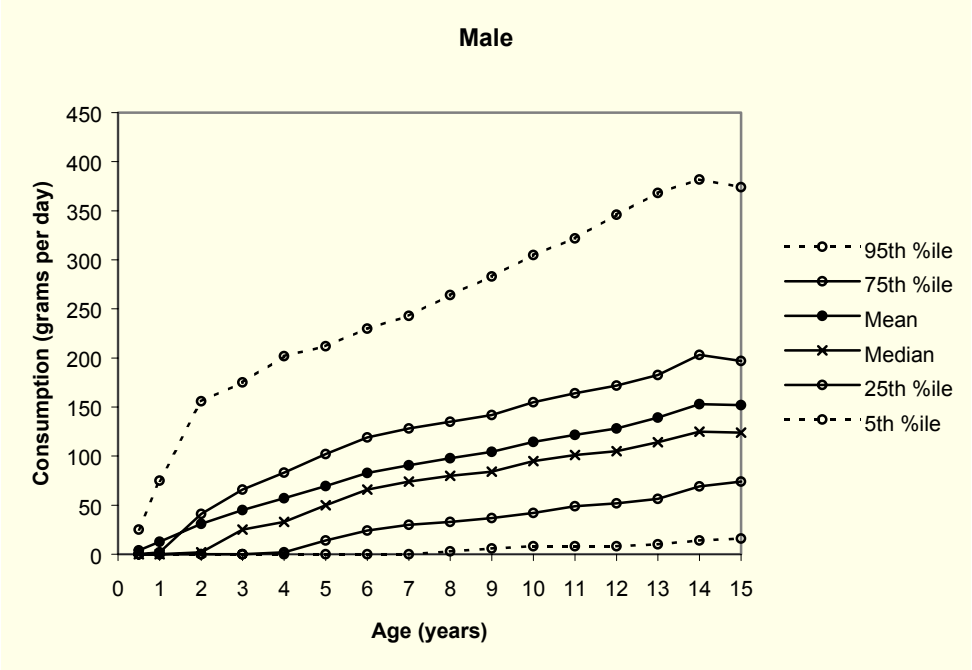
A.6.f. Vegetables

Consumption of green and leafy vegetables was reported for 1693 (85.6%) of the 1979 living evaluable in-area participants whose CATI data were used for dose estimation (855 women, 838 men). For 154 of these 1693 (74 women, 80 men), the CATI respondent was unable to provide estimates of the quantities consumed. Figure IX.A-6 summarizes the distributions of vegetable consumption by sex and age. Each participant's consumption, expressed as grams per day, was calculated from the information reported by the CATI respondent.

As shown in Figure IX.A-6, the reported consumption of vegetables increased steadily with age for both sexes, to about 130 grams per day for women and over 150 grams per day for men by the teenage years. For both sexes, the proportion who were reported not to consume vegetables decreased from over 90% at 6 months of age to 9% at 6 years of age. The proportions of nonconsumers continued to decrease slightly at older ages, ranging between 2% and 4% after age 9.

Figure IX.A-6. Vegetable Consumption, by Sex and Age



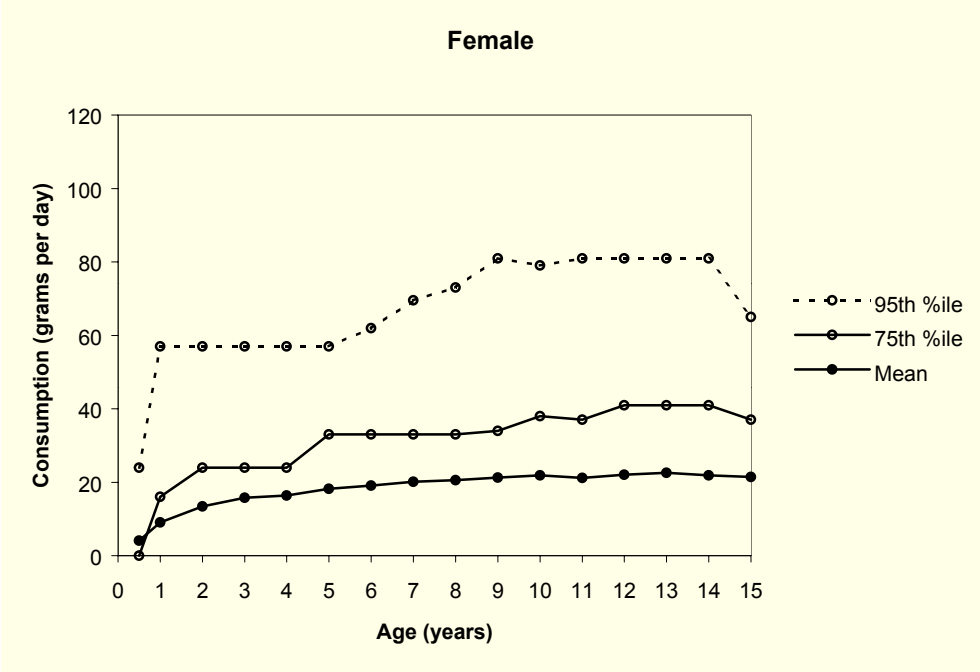


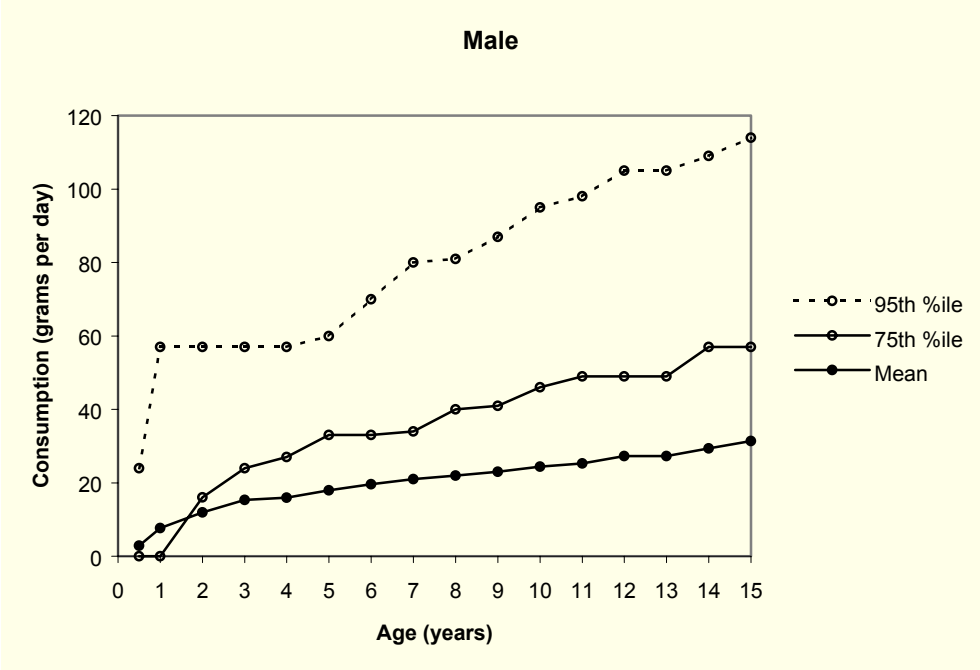
A.6.g. Free Range Chicken Eggs

Consumption of free range chickens was reported for 1057 (53.4%) of the 1979 living evaluable in-area participants whose CATI data were used for dose estimation (552 women, 505 men). For 64 of these 1057 (34 women, 30 men), the CATI respondent was unable to provide estimates of the quantities consumed. Figure IX.A-7 summarizes the distributions of free range chicken egg consumption by sex and age. Each participant's consumption, expressed as grams per day, was calculated from the information reported by the CATI respondent. At every age shown in the figures, fewer than half of the participants were reported to consume free range chicken eggs, and the 5th and 25th percentiles and medians were consequently all zero; therefore these three statistics are omitted from the figure for clarity.

As shown in Figure IX.A-7, the reported consumption of free range chicken eggs increased steadily with age for men, but less so for women. For women, consumption increased to about 16 grams per day by age 3, then increased only slowly to about 22 grams per day by the teenage years. For men consumption increased to about 15 grams per day by age 3, and then continued to increase to about 30 grams per day by the teenage years. For both sexes, the proportion who were reported not to consume free range chicken eggs decreased from over 85% at 6 months of age to just over 50% at 5 years of age, and remained between 50% and 57% for all older ages.

Figure IX.A-7. Free Range Chicken Egg Consumption, by Sex and Age





A.7. Age Distribution for the Alternative Representations to Exposure

As described in section VIII.B.3.b, two alternative representations of exposure to Hanford's ^{131}I were defined, in order to assess whether there might be evidence of a radiation effect that was not apparent from the dose-response analyses using the individual dose estimates calculated by the CIDER program. These alternative representations of exposure were categorical variables, specifically the geostratum and a dichotomous variable defined to identify participants likely to have relatively high versus relatively low exposures (see section VIII.B.3.b.2). Since 1) the definitions of these alternative representations were based entirely (geostratum) or partially (dichotomous exposure variable) on geostratum, and 2) both years of birth and years during which HTDS examinations were performed varied by geostratum, it was of interest to examine whether the participants' ages at HTDS examination were correlated with either alternative representation of dose. Table IX.A-8 shows that the age at HTDS examination varies somewhat by geostratum, with mean age ranging from 49 years for the Okanogan and Ferry/Stevens County geostrata, to 53 years for the Adams County geostratum. This reflects the fact that participants in the Okanogan and Ferry/Stevens geostrata were selected only for the Pilot Study phase, while those from Adams County were selected only during the Full Study phase. In addition, the birth years from which participants were selected were 1942-1946 for the Okanogan and Ferry/Stevens geostrata, while for Adams County they were from 1940-1945.

Table IX.A-8. Age at HTDS Examination by Geostratum

Geostratum	No.	Median	Minimum	Maximum	Mean	St. Dev.
Richland	352	51	46	53	50	1.6
Pasco/Kennewick	1009	52	46	57	52	2.1
Walla Walla City	264	50	46	55	50	1.9
Benton Co.	734	52	46	57	52	2.4
Franklin Co.	149	50	45	57	51	3.1
Walla Walla Co.	334	50	46	55	50	1.9
Okanogan Co.	139	49	46	55	49	2.0
Ferry/Stevens Cos.	138	49	46	54	49	1.5
Adams Co.	321	53	50	57	53	1.6

As can be seen in Table IX.A-9, the age at HTDS examination also differed slightly by the dichotomous exposure variable, with the mean age at HTDS examination 2 years higher in the high exposure group compared to the low exposure group.

Table IX.A-9. Age at HTDS Examination by Dichotomous Exposure Variable

Exposure Group	No.	Median	Minimum	Maximum	Mean	St. Dev.
Low	677	50	46	57	50	2.7
High	580	52	47	57	52	2.0

Although the two tables above indicate that the geostrata, as well as the high and low exposure groups, differed somewhat with respect to the distributions of age at HTDS examination, the differences are rather small, with a maximum difference of 4 years in average age. Although cumulative incidence of disease outcomes or prevalence of thyroid UDAs is likely to increase with age, differences of only a few years of age are unlikely to cause large increases. Nevertheless, the analyses of disease outcomes and thyroid UDAs in relation to these two alternative representations of exposure were adjusted for age at HTDS examination.

B. Estimated Radiation Doses to the Thyroid from Hanford ^{131}I

As described in section VI above, estimates of thyroid radiation doses from atmospheric releases of Hanford's ^{131}I were calculated using the computer program CIDER, which was developed by the HEDR Project (137). Specifically, CIDER calculated estimates of doses received by an individual during any times from December 1944 to the end of 1957, that he or she reports being inside the 246-by-306 mile HEDR geographical domain (Figure II.A-1).

It is important to understand that CIDER does not calculate any contribution to a person's dose for periods he or she reports being outside the HEDR domain. This does not reflect an assumption that persons were not exposed while outside the domain, but rather the difficulty of accurately estimating doses received at long distances from Hanford, and the likelihood that such doses were small. A fundamental objective of the HEDR project in determining the domain's boundaries was to ensure a high likelihood that individuals could not receive appreciable doses while outside the domain. In particular, the domain was defined by the HEDR Project to include as much as possible of the region over which appreciable doses might have been received, while taking into account the decreasing reliability of ^{131}I atmospheric transport modeling at longer and longer distances.

Based on the residence histories obtained from the CATIs and Expanded In-Person Interviews, 3191 (93%) of the 3440 living evaluable participants lived within the HEDR domain at least some time from December 1944 to the end of 1957. These are the participants for whom CIDER can compute a dose estimate. For convenience, these 3191 participants are designated as "in-area" participants in this report (see Table IX.B-1 below). The residence histories of the remaining 249 living evaluable participants (7% of the 3440) included no residence within the HEDR domain from December 1944 to the end of 1957. These 249 individuals are designated as "out-of-area" participants.

Dosimetric data were obtained from CATIs for 2123 (62%) of the 3440 living evaluable participants, and from Expanded-IPIs for the remaining 1317 (38%).

Table IX.B-1. Summary of Dosimetry Interview Types and In-Area Status of 3440 Living Evaluable Participants

	In-Area Status		
Type of Dosimetry Interview	In-Area	Out-of-Area	Total
CATI	1979	144	2123
Expanded In-Person	1212	105	1317
Total	3191	249	3440

B.1. Calculation of Estimated Thyroid Radiation Doses for In-Area Participants

Three sets of dose estimates were calculated for the 3191 in-area participants. The sources of data used for these three sets of doses are summarized in Table IX.B-2.

Table IX.B-2. Characteristics of Primary and Alternative Sets of Radiation Dose Estimates

	Primary Dose Estimates	First Alternative Dose Estimates	Second Alternative Dose Estimates
Use CATI specifics regarding amounts and sources of food and milk?	Yes	No	Yes
Source of default values	HEDR	HEDR	HTDS*

* Note that for expanded IPIs HEDR defaults were used for consumption other than milk.

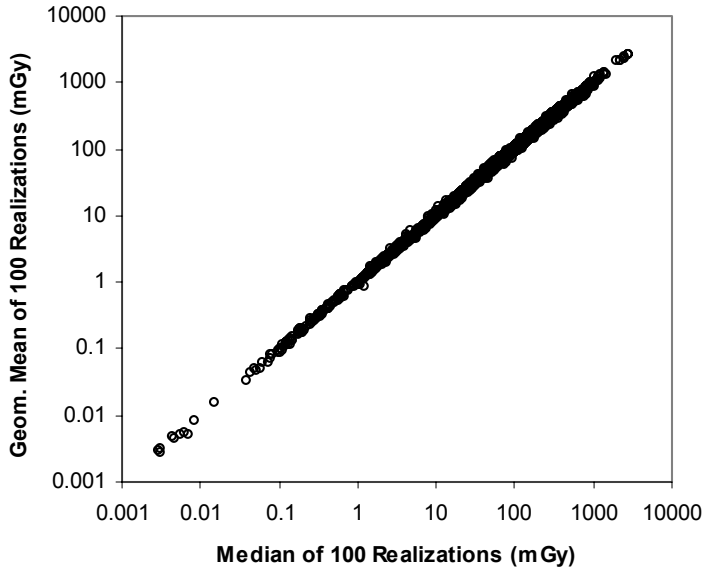
In the dose-response analyses reported below, the primary emphasis is given to results based on the primary set of dose estimates. Analyses using the alternative sets of dose were performed primarily to assess the sensitivity of the dose-response results to the type of dose estimate.

B.2. Point Estimates and Uncertainty of Doses

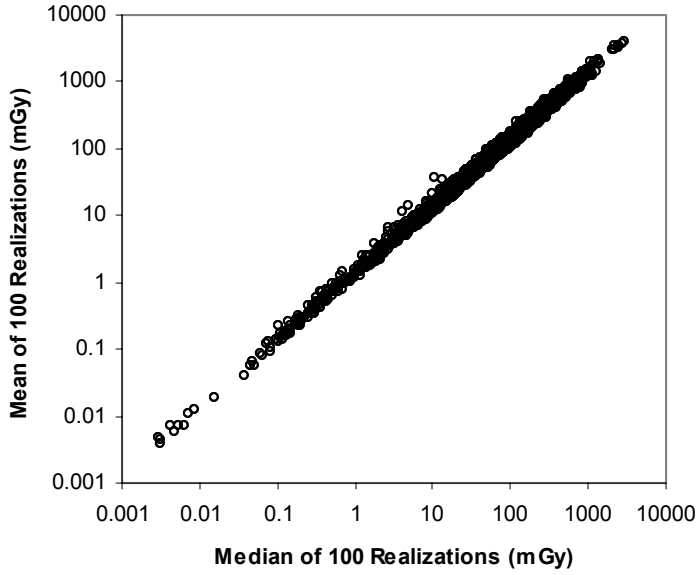
As described in section VIII.B.3.a above, the CIDER program actually returns 100 estimates or "realizations" of each participant's dose. For many purposes, in particular for the conventional analyses of dose-responses described in section VIII.B.3.a above, it is important to have a single number or "point estimate" to serve as each participant's estimated dose. Three obvious candidates for the point estimate are the median, mean, and geometric mean of the 100 realizations. These three point estimates were calculated for each of the 3191 in-area living evaluable participants. It was expected that the three point estimates would be highly correlated with each other, and this is confirmed in Figure IX.B-1 below.

Figure IX.B-1. Scatterplots of Geometric Mean and Mean Doses versus Median Dose

A. Geometric Mean versus Median



B. Mean versus Median

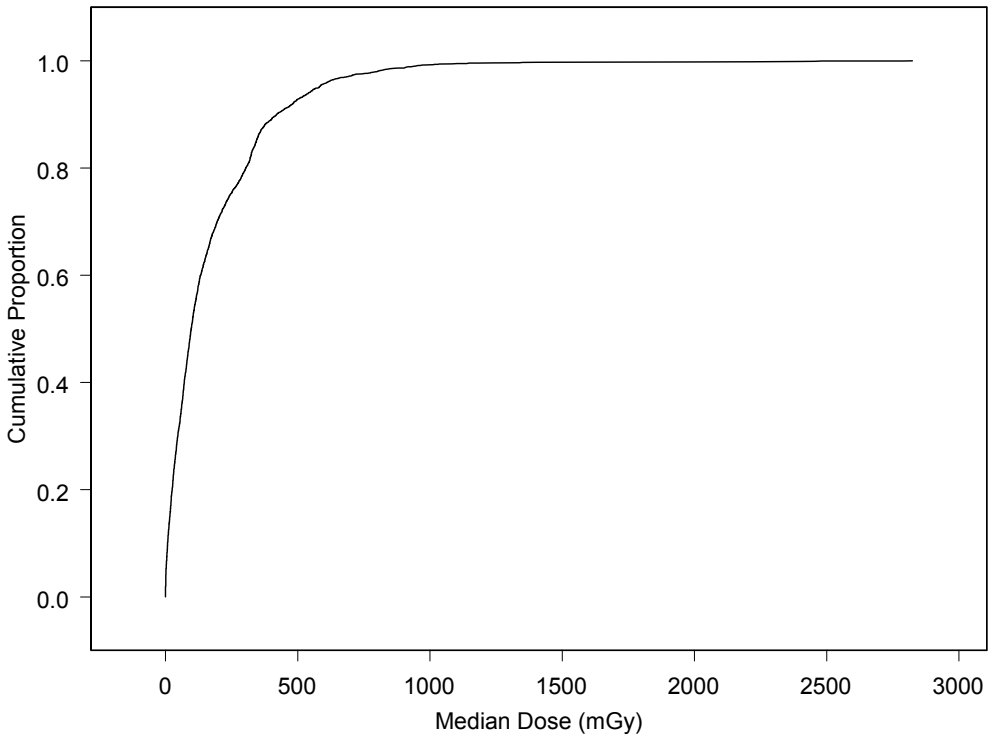


The distributions of each participant's 100 dose estimates tended to be roughly lognormally distributed. Therefore the medians and geometric means were nearly equal for most participants, as can be seen in Panel A of Figure IX.B-1. Furthermore, due to the approximate lognormality of each participant's 100 realizations, the mean doses tended to be somewhat larger than the medians (Figure IX.B-1, Panel B).

Because of the very high degree of correlation among the three possible point estimates, it can be expected that they will give very similar results in the analyses of radiation dose-responses, at least in terms of statistical significance. Therefore the remainder of this report focuses primarily on the median as the point estimate of participants' doses. For simplicity, the terms "doses" or "estimated doses" will refer to the median dose estimates unless otherwise indicated.

Figure IX.B-2 displays the cumulative frequency distribution (CDF) of the median doses. The shape of the CDF indicates that the distribution of median doses is strongly skewed to the right. The majority of participants have relatively low doses, while the rest have doses that are spread over a wide range of higher values.

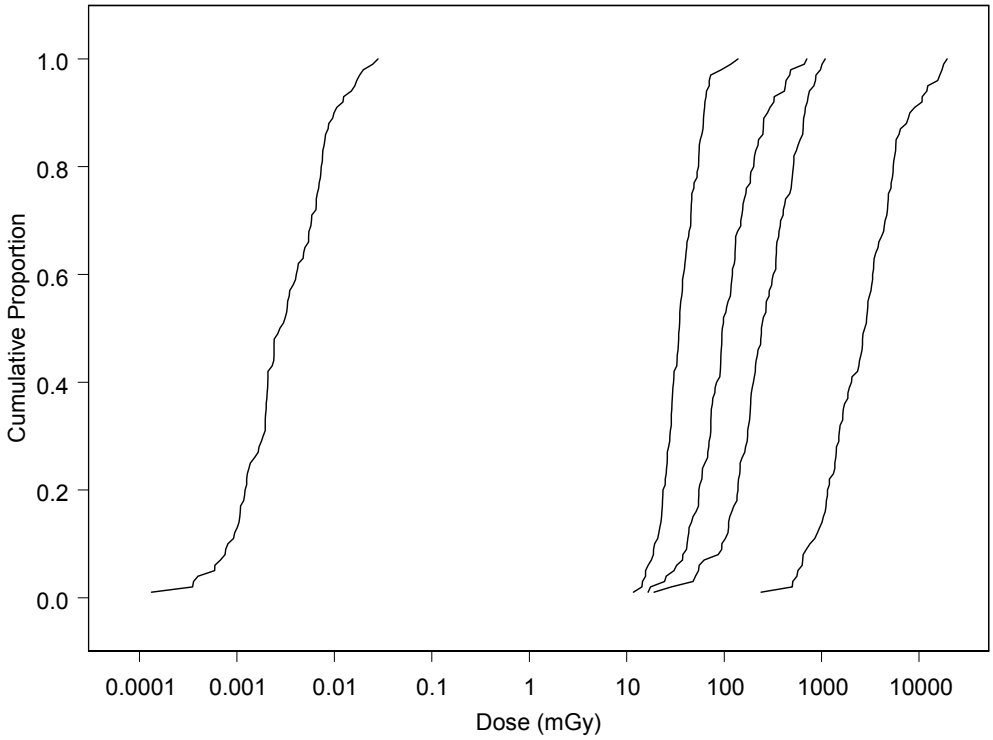
Figure IX.B-2. Cumulative Distribution of Median Dose Estimates for 3191 In-Area Living Evaluable Participants



The uncertainty in each participant's dose estimate is represented by the variation among his or her 100 dose estimates provided by CIDER (see section VIII.B.3.a above). This is illustrated in Figure IX.B-3. Empirical cumulative distribution functions of the 100 dose realizations for each of five selected participants are shown in the figure. The five participants were chosen on the basis of their estimated doses (i.e., the medians of their 100 dose realizations) to cover the entire range of dose estimates. Specifically, the participants in the figure are those with (from left to right) the smallest dose, the 25th, 50th, and 75th

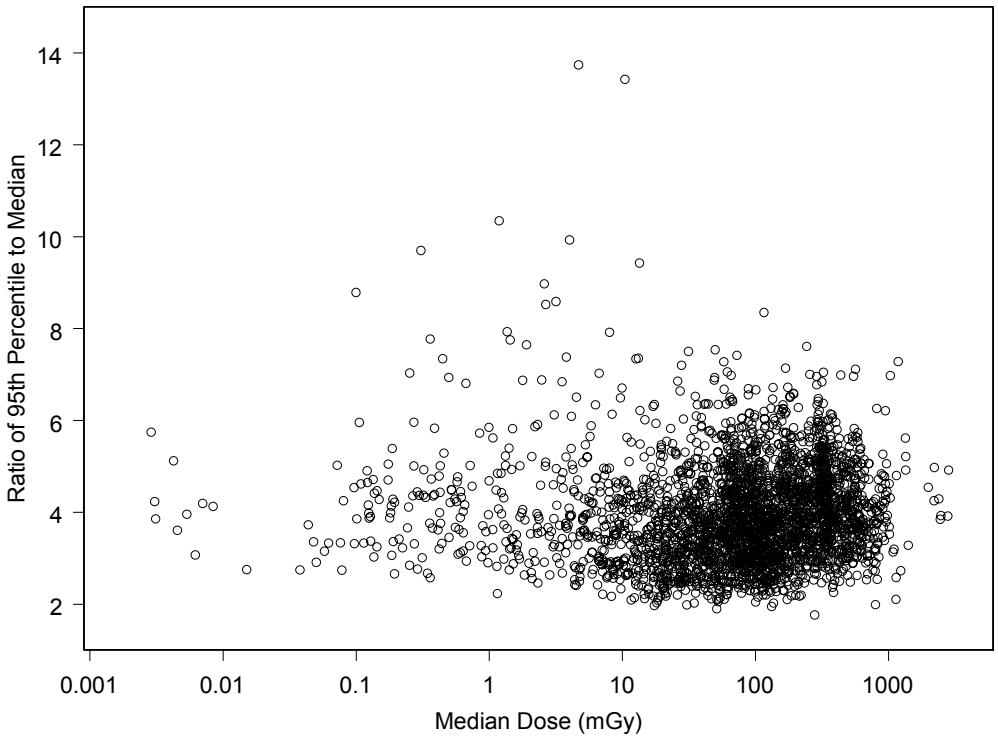
percentile doses, and the largest dose among all 3191 living evaluable in-area participants. It is evident from Figure IX.B-3 that the distributions of 100 realizations are approximately normally distributed. Moreover, the fact that the curves are approximately parallel suggests that the variances of log-transformed dose realizations, or equivalently, the geometric standard deviations of the dose realizations, are roughly the same for each participant.

Figure IX.B-3. Cumulative Distribution of 100 Dose Estimates for Five Selected Participants



One simple and useful characterization of the magnitude of the uncertainty is the ratio of an upper percentile of the 100 realizations, e.g., the 95th percentile, to the median. Among the 3191 in-area living evaluable participants, these ratios had a median value of 3.8 and ranged from 1.8 to 13.7, although the ratio was less than 10.0 for all but three of the 3191 participants. Only 10% of the ratios were less than 2.7, and only another 10% were greater than 5.3. Figure IX.B-4 displays these ratios in relation to the median dose estimates.

Figure IX.B-4. Scatterplot of Ratio of 95th Percentile to Median versus Median Dose



The clustering of the ratios around a value of about 4 is evident in Figure IX.B-4, and there is no indication that the magnitude of uncertainty varies in relation to the median dose.

Another simple characterization of the magnitude of the uncertainties of the estimated doses is the geometric standard deviation or GSD (see section VIII.B.3.a above). For the 3191 living evaluable in-area participants, i.e., the study participants for whom the CIDER program was able to calculate dose estimates, the GSDs ranged from 1.56 to 5.42, with a mean of 2.18.

A total of 890 living evaluable in-area participants (28%) had dose estimates with GSDs less than 2.0. In its review of the HTDS Draft Final Report, the National Research Council (NRC) questioned how GSDs less than 2 could occur, reasoning as follows (138). In CIDER's dose calculations, dose conversion factors (DCFs) were treated as uncertain parameters. For example, CIDER uses age-specific ingestion DCFs to convert total ingested Curies of ^{131}I activity (from milk, food, etc.) into dose (measured in mGy). Similarly, CIDER uses age-specific inhalation DCFs to convert inhaled ^{131}I activity to dose. In CIDER, the DCFs for the ingestion and inhalation pathways, which accounted for most of the dose received by the majority of study participants, were assumed to be lognormally distributed with a GSD of 2.0 (10509). Since the GSD of the product of two uncertain variables is the product of their respective GSDs, the NRC reasoned, the GSDs of the doses should rarely if ever be less than 2. The NRC further noted that only GSDs of 2 or greater were reported for representative dose calculations performed by the HEDR project [10493].

The NRC failed to recognize that dose estimates with GSDs less than 2.0 were a predictable consequence of the fact that CIDER's calculation of doses involves addition of doses after activity levels are multiplied by DCFs. Specifically, the last step of CIDER's dose calculation is the addition of estimated doses from multiple pathways (ingestion, inhalation, and immersion) and time periods defined by age

and/or calendar year. Each of the estimated dose components in this addition has its own uncertainty, with GSD 2 or greater for ingestion or inhalation components. Now consider the addition of two lognormally distributed (or approximately lognormally distributed) variables with similar GSDs. If the first variable has a much larger geometric mean than the second, the GSD of the sum will generally be close in value to the GSD of the first variable. However if the two variables have similar geometric means, the GSD of their sum will be substantially less than either variable's GSD. Both of these situations occurred in the calculation of dose estimates for the various HTDS participants. Many participants received a large majority of their accumulated dose from one pathway (e.g., ingestion) and one time period (e.g., all or part of 1945). For such participants, the GSD of the total dose was therefore close to the GSD of that dominant component, i.e., 2 or greater. However, for other participants there were two or more components of dose having roughly similar geometric means. When added together, these produced total doses with GSDs less than 2. The NRC also failed to recognize that the representative dose calculations reported by the HEDR project were not informative in this regard. For example, consider the representative doses reported in Table 1 of the paper by Farris et al. [10493]. The estimated doses for the hypothetical people represented in that table were dominated by the component accumulated through ingestion during 1945. Consequently the GSDs for all of the examples in the table were 2 or greater.

B.3. Distributions of Primary Dose Estimates

The primary estimates of radiation dose for the 3191 in-area living evaluable participants ranged from a minimum of 0.0029 mGy to a maximum of 2823 mGy, with a median of 97 mGy. The mean and standard deviation of the distribution of estimated doses were 174 mGy and 224 mGy, respectively. The distribution of dose estimates was quite heavily skewed, as shown in Figure IX.B-2 above. As shown in Table IX.B-3 below, the distributions of median doses did not differ markedly between women and men.

Table IX.B-3. Frequency Distribution of Estimated Thyroid Radiation Dose, by Sex

Estimated Thyroid Radiation Dose (mGy)	Living Evaluable Participants					
	Female		Male		Total	
	No.	%	No.	%	No.	%
< 10	182	11.2	186	11.9	368	11.5
10-49	320	19.7	314	20.0	634	19.9
50-99	313	19.3	310	19.8	623	19.5
100-149	220	13.6	171	10.9	391	12.3
150-199	126	7.8	109	6.9	235	7.4
200-299	139	8.6	148	9.4	287	9.0
300-399	144	8.9	160	10.2	304	9.5
400-999	171	10.5	154	9.8	325	10.2
1000+	7	0.4	17	1.1	24	0.8
Total In-Area	1622	100	1569	100	3191	100
Out of Area	125	7.2	124	7.3	249	7.2
Total	1747	100	1693	100	3440	100

Twenty-four (0.8%) of the 3191 in-area living evaluable participants had dose estimates greater than 1000 mGy, and only seven (0.2%) had estimates over 2000 mGy. Summary statistics for the distributions of estimated doses are shown by geostratum in Table IX.B-4. As expected, the estimated doses tended to be higher for participants in the Richland, Pasco/Kennewick, and Benton, Franklin and Adams County geostrata. They tended to be lowest for the Okanogan and Ferry/Stevens County geostrata, and intermediate for the two Walla Walla geostrata.

Table IX.B-4. Summary of Estimated Radiation Doses (in mGy) to the Thyroid from Hanford ^{131}I, by Geostratum

Geostratum	No.	Median	Minimum	Maximum	Mean	St. Dev.
Richland	348	101	0.1	2455	220	284
Pasco/Kennewick	910	242	.003	1235	255	200
Walla Walla City	250	64	.06	745	74	76
Benton Co.	656	83	.05	2823	170	311
Franklin Co.	141	173	.004	1028	234	215
Walla Walla Co.	320	66	.1	1016	83	93
Okanogan Co.	125	5	.003	158	11	19
Ferry/Stevens Cos.	131	32	.02	128	36	28
Adams Co.	310	161	.008	584	166	101
Total	3191	97	.0029	2823	174	224

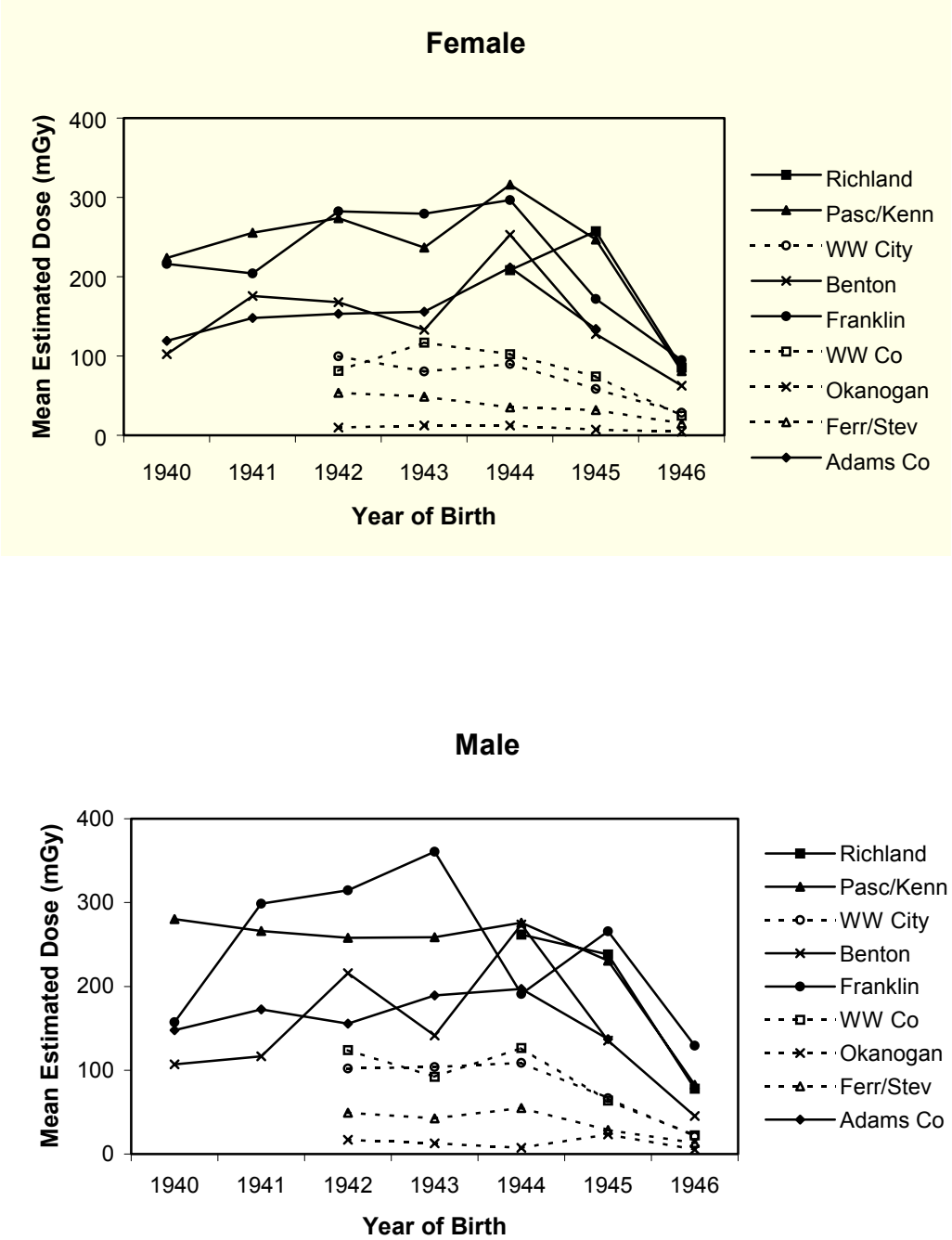
Summary statistics for the distributions of estimated doses are shown by sex and birth year in Table IX.B-5 below. The distributions of estimated doses were similar for men and women. The arithmetic mean doses are slightly larger for men (177 mGy) than women (171 mGy), but the medians are quite similar (96 mGy for men, 99 mGy for women). Seventeen (1.1%) of the 1569 men and 7 (0.4%) of the 1622 women had doses above 1000 mGy. Four of the seven participants with doses over 2000 were female.

Table IX.B-5. Summary of Estimated Radiation Doses (in mGy) to the Thyroid from Hanford ^{131}I, by Sex and Year of Birth

Sex	Year	No.	Median	Minimum	Maximum	Mean	St. Dev.
Female	1940	111	122	.04	547	160	132
	1941	133	164	.003	935	202	170
	1942	224	102	.003	952	166	175
	1943	236	102	.007	652	152	146
	1944	416	125	0.05	2823	230	329
	1945	318	94	0.1	956	171	187
	1946	184	29	0.3	373	50	58
	1940-46	1622	99	.003	2823	171	220
Male	1940	107	163	.05	1102	199	187
	1941	122	165	0.3	782	204	163
	1942	211	107	.005	1016	174	166
	1943	248	118	0.2	1235	179	177
	1944	413	114	.003	2455	228	324
	1945	290	73	.08	975	166	208
	1946	178	28	.006	717	48	78
	1940-46	1569	96	.003	2455	177	228
Total	1940-46	3191	97	.0029	2823	174	224

Doses tended to be higher for participants born in 1940 – 1941, to drop somewhat for those born in 1942 – 1943, then to increase again for those born in 1944 . This pattern was largely an artifact of the way in which the study cohort was defined. As described above, the 1940 and 1941 birth cohorts were limited to Benton, Franklin and Adams counties (including Pasco and Kennewick), since these counties were expected to provide participants with relatively high doses. The other counties, from which participants would have been expected to have lower doses, were not included in the 1940 and 1941 cohorts. The effect of this exclusion is shown in Figure IX.B-5 below, which displays the mean estimated doses by sex, birth year, and geostratum.

Figure IX.B-5. Mean of Estimated Median Thyroid Radiation Dose (in mGy) from Hanford ^{131}I by Sex, Year of Birth, and Geostratum



Twenty-three of the 24 participants with estimated doses greater than 1000 mGy were in the Richland, Pasco/Kennewick, Benton or Franklin County geostrata (one was in Walla Walla County). All but three of the 24 were born in 1944, as were all of the seven with estimated doses over 2000 mGy.

Participants' ages at the time of their HTDS examinations ranged from 45 to 57 years. As shown in Table IX.B-6, the participants who were youngest when examined tended to have lower doses, since many of them were born too late to be exposed during the period of highest releases in early and mid 1945.

Table IX.B-6. Summary of Radiation Doses (in mGy) to the Thyroid from Hanford ^{131}I by Sex and Age at HTDS Examination

Sex	Age at Exam	No.	Median	Minimum	Maximum	Mean	St. Dev.
Female	45-46*	33	21	0.3	105	34	29
	47	84	35	0.3	472	69	88
	48	136	50	0.1	946	96	137
	49	165	69	0.1	1028	114	147
	50	182	103	0.06	2823	205	320
	51	359	126	0.003	1349	210	220
	52	256	110	0.1	2792	211	296
	53	112	118	0.007	821	169	158
	54	105	163	2.1	676	199	160
	55	133	126	0.004	935	177	161
	56	54	124	0.003	450	146	122
	57	3	128	56	195	126	70
Male	46	26	25	5	717	61	137
	47	106	33	0.006	486	60	79
	48	126	43	0.003	931	104	170
	49	146	69	0.1	1083	138	198
	50	186	80	0.07	1015	168	215
	51	341	120	0.005	2455	220	294
	52	253	120	0.1	1989	214	262
	53	113	171	0.3	1337	216	198
	54	111	168	0.005	782	191	145
	55	107	163	1.4	793	205	179
	56	47	123	0.3	1102	169	188
	57	7	196	.05	368	193	144

* Only one person was 45 years old at the time of examination.

There were two major differences in the dose calculations for participants with CATIs versus those with Expanded In-Person Interviews:

- The first, and perhaps most obvious difference, was the potential availability from CATIs of specific, detailed information about quantities and sources of the milk and other food products consumed by the participant during 1944 – 1957. The CIDER program provided default estimates of these characteristics whenever they were not specified by HTDS. Thus the CIDER defaults were used for all participants with dose calculated from Expanded In-Person Interview. For those with CATI dosimetry data, however, the CIDER defaults were used only when necessary, i.e., when the CATI respondent was unable to provide the information.

- The second major difference between doses calculated from CATI and Expanded In-Person Interview data concerned the contributions to participants' doses from breastfeeding. The CATI included

questions regarding whether or not the participant was breastfed any time after the start of ^{131}I releases from Hanford. However, it was anticipated that many, if not most participants, without CATI respondents would be unable to answer such a question accurately. Therefore the Expanded In-Person Interview did not include questions regarding breastfeeding of the participant. Similarly, since participants without CATI respondents could not be expected to recall details of early life such as the age at which they began drinking cow's milk, no such questions were included in the Expanded In-Person Interview. In the absence of data on these characteristics, the CIDER model assumed that cow's milk consumption began at birth. Therefore all 1212 in-area living evaluable participants with an Expanded In-Person Interview were effectively assumed to have begun drinking cow's milk at birth.

The impact of interview type on the estimated doses is shown in Table IX.B-7.

Table IX.B-7. Summary of Estimated Radiation Doses (in mGy) to the Thyroid from Hanford ^{131}I, by Type of Dosimetry Interview and Year of Birth

Interview	Birth Year	No.	Median	Minimum	Maximum	Mean	St. Dev.
CATI	1940	116	150	0.05	1102	209	191
	1941	115	139	1.7	935	208	204
	1942	277	102	0.003	953	173	181
	1943	284	106	0.007	1235	170	184
	1944	511	91	0.003	1143	151	195
	1945	408	59	0.08	943	105	134
	1946	268	29	0.006	717	51	74
	1940-46	1979	81	0.003	1235	140	174
Expanded IPI	1940	102	125	0.04	707	145	112
	1941	139	187	0.003	524	198	128
	1942	158	104	0.06	1016	164	149
	1943	200	106	0.1	498	160	129
	1944	319	247	0.1	2823	355	436
	1945	200	269	0.5	975	299	237
	1946	94	28	1.2	236	46	48
	1940-46	1212	154	0.003	2823	229	279

The average doses (i.e., the means or medians in Table IX.B-7) are generally similar for each year except 1944 and 1945. In those two years, however, the doses based on Expanded In-Person Interviews are notably larger (arithmetic means 355 and 299 mGy, respectively) compared to those based on CATI input data (151 and 105 mGy). This difference reflects the assumption that participants without CATI dosimetry data were assumed to drink cow's milk from birth. As described in the paragraphs above, this likely led to overestimation of the doses for some of the participants with doses based on Expanded In-Person Interviews who were in fact breastfed. In addition, CATI respondents reported that the majority of participants did not consume fresh cow's or goat's milk or milk products in the first months of life (e.g., 69% at 6 months of age; see section IX.A.6.d above).

Table IX.B-8 displays the dose distributions according to age at first exposure to ^{131}I from Hanford, age at HTDS examination, and estimated thyroid dose from the Nevada Test Site (NTS). Participants with prenatal exposure have rather lower doses than other participants, in part since nearly all of the 1946 birth stratum, which missed the months of highest ^{131}I releases from Hanford, were exposed in utero. Participants who were first exposed to ^{131}I from Hanford before 180 days of age also have somewhat lower doses for a similar reason. Participants who were ≤50 years old at the time of their HTDS examinations also had somewhat lower doses, since they tended to be in the later birth year strata. Participants with relatively higher estimated thyroid doses from the NTS tended to have lower doses from

Hanford, in part due to residence. There were no major differences in the doses of those who had a history of any cancer other than thyroid compared to those with no such history.

Table IX.B-8. Summary of Estimated Radiation Doses (in mGy) to the Thyroid from Hanford ^{131}I, by Age at Exposure and HTDS Examination, Estimated Thyroid Dose from NTS and History of Any Cancer Other Than Thyroid

Covariate		No.	Median	Minimum	Maximum	Mean	St. Dev.
Prenatal exposure?	Yes	1034	58	.038	2206	135	194
	No	2157	118	.003	2823	193	235
1st exposure before age 180 days?	Yes	1478	75	.038	2823	172	269
	No	1713	115	.003	1350	176	176
Age at HTDS Examination > 50?	Yes	2001	128	.003	2792	203	233
	No	1190	61	.003	2823	125	199
NTS thyroid dose > 5.3 mGy?	Yes	1567	66	.003	2792	128	206
	No	1622	145	.003	2823	218	232
History of any cancer other than thyroid?	Yes	248	104	.003	2823	194	310
	No	2938	96	.003	2792	172	215

Table IX.B-9 displays distributions of estimated dose in relation to participants' histories of various types of medical and dental radiation exposures. The thyroid doses from Hanford do not differ greatly according to the presence or absence of the various kinds of exposure.

Table IX.B-9. Summary of Estimated Radiation Doses (in mGy) to the Thyroid from Hanford ^{131}I, by Medical and Dental Radiation History

Have You Ever Had:		No.	Median	Minimum	Maximum	Mean	St. Dev.
CAT scan of the upper body?	Yes	775	98	.003	2823	174	188
	No	2374	96	.003	930	174	236
Diagnostic x-rays of the head?	Yes	1191	90	.003	2482	164	215
	No	1964	102	.003	2823	179	228
Diagnostic x-rays of the neck?	Yes	966	112	.003	2823	195	257
	No	2201	90	.003	2455	164	207
Diagnostic x-rays of the chest or upper? body, including mammograms?	Yes	2821	96	.003	2823	176	228
	No	352	99	.005	1410	161	191
Diag. x-rays of the stomach or mid-back?	Yes	692	96	.003	2482	165	211
	No	2428	97	.003	2823	176	227
Barium enema?	Yes	825	94	.003	2823	174	223
	No	2334	99	.004	2792	174	225
Upper GI?	Yes	1146	99	.003	2823	181	226
	No	2031	96	.004	2792	170	223
Intravenous pyelogram or IVP?	Yes	398	100	.003	1337	185	215
	No	2759	97	.003	2823	172	225
Fluoroscopy of the upper body?	Yes	246	105	.210	1028	192	222
	No	2915	96	.003	2823	172	224
Other nuclear scan?	Yes	217	92	.122	1337	185	219
	No	2945	97	.003	2823	173	225
Radiation treatment for any cancer other than thyroid?	Yes	39	119	.413	1349	202	275
	No	3147	97	.003	2823	174	223
Dental x-rays that did not usually include a lead shield over the neck area?	Yes	1648	95	.003	2482	170	222
	No	1543	99	.005		178	226

Table IX.B-10 displays distributions of estimated dose in relation to participants' occupational histories. The 371 living evaluable in-area participants who reported ever working in a nuclear facility had somewhat higher estimated thyroid doses from Hanford (median 148 mGy).

Table IX.B-10. Summary of Estimated Radiation Doses (in mGy) to the Thyroid from Hanford ^{131}I, by Occupational History

Have You Ever Worked in Any of the Following:		No.	Median	Minimum	Maximum	Mean	St. Dev.
Any metal industry?	Yes	238	85	.003	1016	176	204
	No	2953	98	.003	2823	174	226
Any nuclear facility?	Yes	371	148	.015	2455	248	280
	No	2820	93	.003	2823	164	214
Any other industry or occupation where you may have been exposed to radioactive materials or x-rays?	Yes	442	92	.003	2823	172	258
	No	2749	98	.003	2792	174	218
Any of the above industries or occupations?	Yes	892	104	.003	2823	196	250
	No	2299	94	.003	2792	166	213

As shown in Table IX.B-11, the distributions of estimated thyroid doses from Hanford did not differ appreciably between living evaluable in-area participants who reported any history of smoking cigarettes, or of any smoking, compared to those without such histories.

Table IX.B-11. Summary of Estimated Radiation Doses (in mGy) to the Thyroid from Hanford ^{131}I, by Smoking History

Have You Ever Smoked Any of the Following:		No.	Median	Minimum	Maximum	Mean	St. Dev.
Cigarettes (unfiltered or filtered)?	Yes	1854	96	.003	2823	177	238
	No	1329	98	.005	2206	169	203
Any of cigarettes, cigar or pipe?	Yes	1900	96	.003	2823	177	237
	No	1283	97	.005	2206	169	204

Since the consumption of milk contaminated with ^{131}I from Hanford was a key source of exposure for many study participants, the relationship between milk and milk product consumption and estimated radiation dose was investigated. Average milk and milk product consumption levels (expressed as the reported average number of 8 oz. servings consumed per day) were calculated as described in section VIII.B.2.c above for each of the 1979 living evaluable participants whose CATI data were used for dose estimation. To calculate each participant's average consumption level, his or her reported total number of 8 oz. servings for a particular type of milk was first calculated by integrating the reported consumption levels over the time periods for which the CATI respondent reported consumption levels of that milk. For example, if a CATI respondent reported that a participant consumed three 8 oz. servings per day over a period of 2 years, the total consumption was $3 \times 2 \times 365 = 2190$ 8 oz. servings for that period. For these calculations participants born in 1946 were assigned milk consumption values of 0 for 1945. Also, participants in the 1940-1945 birth strata who never lived inside the HEDR domain during 1945 were assigned consumption levels of 0 for 1945. If the consumption level for a particular type of milk was

unknown for any or all of the time period in question (because the CATI respondent could not report the quantity of glasses consumed), the total consumption was considered unknown for these calculations.

Two measures of average consumption were calculated. The first measure of average consumption, designated "Average No. of 8 oz. Servings per Day," was obtained by dividing the reported total number of 8 oz. servings for a given time period by the duration of that period in days (e.g., by 365 for average consumption during 1945). The second measure, designated "Average No. of 8 oz. Servings per In Area Day," used a different divisor: the number of days during the period for which (1) the participant lived within the HEDR domain and (2) the level of milk consumption was reported in the participant's CATI.

Average consumption levels were calculated for two time periods: (1) 1945, the year in which by far the largest amount of ^{131}I was released from Hanford (see section II above), and (2) the entire period 1944-1957. Table IX.B-12 summarizes milk consumption data reported for the 1979 participants whose CATI data were used for dose estimation are shown for three types of milk: raw ("backyard") cow's milk, processed cow's milk, and goat's milk, as well as for total cow's milk and total milk (cow's plus goat's). Among the 1979 in-area participants whose CATI data were used for dose estimation, the proportions for whom reported average consumption levels were known exceeded 90% for all types of milk. The proportion for whom the reported average consumption level was zero varied widely according to the type of milk or milk product and time period. For example, the proportion with no consumption of cow's milk (raw or processed) was 37% for 1945, but only 8% for all years. This difference reflects the experience of children who were too young to consume cow's milk in 1945, but did consume it at older ages, and the assignment of 0 consumption in 1945 for participants born in 1946. Since very few participants were reported to have consumed goat's milk or milk products, the median consumption levels for these were zero. Similarly the median consumption levels of raw cow's milk and milk products were all 0, since fewer than half of the participants were reported to have consumed these.

Table IX.B-12. Milk and Milk Product Consumption Levels Reported by CATI Respondents: Distributions and Correlation with Estimated Dose

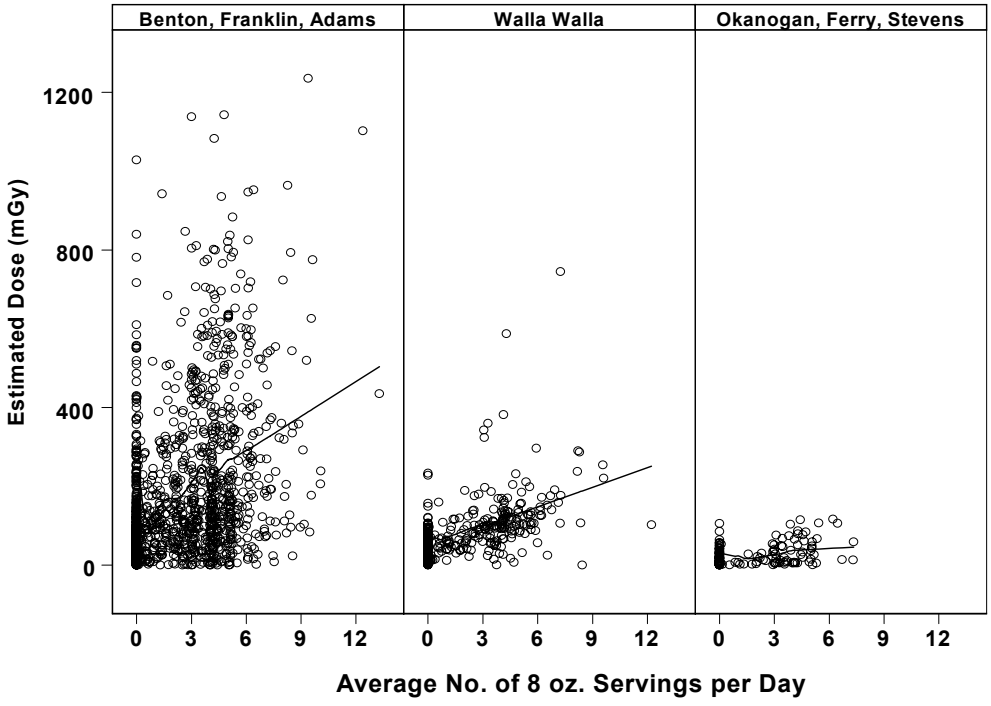
		Participants with Milk and Milk Product Consumption Levels Reported in CATI			Reported Average No. of 8 oz. Servings per Day				Reported Average No. of 8 oz. Servings per In Area Day			
Period	Type of Milk or Milk Product	No.	%	Number (%) with 0	Mean	Median	Range	Corr. with Estimated Dose *	Mean	Median	Range	Corr. with Estimated Dose *
1945	Raw cow's	1927	97%	1351 (70%)	1.05	0	0 – 12.22	0.31	1.09	0	0 – 12.22	0.30
Only	Processed cow's	1904	96%	1016 (53%)	1.07	0	0 – 12.38	0.37	1.19	0	0 – 12.38	0.34
	Total cow's	1866	94%	683 (37%)	2.13	1.37	0 – 12.38	0.57	2.30	1.86	0 – 13.29	0.53
	Goat's	1979	100%	1947 (98%)	0.03	0	0 – 5.69	0.09	0.03	0	0 – 6.00	0.09
	Total	1866	94%	679 (36%)	2.16	1.40	0 – 12.38	0.57	2.33	2.00	0 – 13.29	0.53
1944	Raw cow's	1902	96%	980 (52%)	1.28	0	0 – 11.49	0.23	1.43	0	0 – 11.49	0.21
– 1957	Processed cow's	1854	94%	238 (13%)	2.12	1.66	0 – 12.08	0.31	2.64	2.58	0 – 12.38	0.21
	Total cow's	1807	91%	147 (8%)	3.43	3.58	0 – 12.08	0.46	4.11	4.18	0 – 12.93	0.38
	Goat's	1979	100%	1933 (98%)	0.01	0	0 – 4.29	0.06	0.02	0	0 – 5.00	0.06
	Total	1807	91%	147 (8%)	3.44	3.59	0 – 15.87	0.46	4.12	4.18	0 – 15.87	0.39

* Spearman rank order correlation coefficient

Table IX.B-12 above also shows the correlation between estimated radiation dose and the measures of average milk and milk product consumption. The Spearman rank order correlation coefficients ranged from 0.21 to 0.57 for the various measures of cow's milk and milk product consumption. In contrast, the correlations were quite low for goat's milk and milk products: since these were consumed by only a small minority of the participants, most of the variability of the estimated doses occurred among the nonconsumers of goat's milk and milk products, resulting in the low correlation.

In view of the large number of other factors that influenced the participants' doses, the magnitude of the correlations between estimated dose and these aggregate measures of cow's or total milk and milk product consumption is noteworthy. The effect of milk and milk product consumption on estimated doses of course depends on other factors, in particular the participant's residence history. For example the HEDR model implies that consuming an average of one 8 oz. servings per day throughout 1945 resulted in a higher dose for residents of Franklin county (immediately east of the Hanford site) than for, say, Jefferson County, Oregon (in the southwest corner of the HEDR domain). Figure IX.B-6 displays the relationship between estimated thyroid dose and one measure of milk and milk product consumption, the reported average number of 8 oz. servings per in area day of the total of cow's and goat's milk for 1945. In order to display, at least approximately, the effect of residence location on the relationship between consumption and estimated dose, the participants were divided into three groups based on geostratum: Benton, Franklin and Adams counties (including Richland, Pasco and Kennewick); Walla Walla County (including Walla Walla City); and Okanogan, Ferry and Stevens Counties. While a participant's geostratum (i.e., county of mother's usual residence at the participant's birth) does not correspond perfectly to his or her residence history, it provides a reasonable approximation.

Figure IX.B-6. Estimated Dose in Relation to Reported Consumption of Cow's and Goat's Milk and Milk Products During 1945, by Geostratum



Note: Richland and Pasco/Kennewick geostrata are included within "Benton, Franklin, Adams" and Walla Walla city geostratum is included within "Walla Walla".

The curves in Figure IX.B-6 show smoothed estimates of the average doses, as function of the consumption level, for the three groups. A trend of increasing dose with increasing consumption is evident for the figure. Moreover consumption had a stronger effect on dose for participants in the Benton, Franklin and Adams county geostrata, compared to the other two groups. For example, based on the fitted curves in Figure IX.B-6, the average estimated dose for participants in the Benton, Franklin and Adams County geostrata increased from 85 mGy for those with no consumption of cow's or goat's milk or milk products, (i.e., zero 8 oz. servings per in area day) to 219 mGy for those with an average consumption of four 8 oz. servings per in area day. In contrast, for the Walla Walla County geostrata, the mean estimated dose increased from 38 mGy to 119 mGy, while for the Okanogan, Ferry, and Stevens County geostrata, the mean estimated dose increased from 15 mGy for nonconsumers to 36 mGy for those with an average consumption of four 8 oz. servings per in area day.

As described in Section VIII.B.3.b, two alternative representations of exposure to Hanford's ^{131}I were defined, to assess whether there might be evidence of a radiation effect that was not apparent from the dose-response analyses using the individual dose estimates calculated by the CIDER program. These alternative representations of exposure were categorical variables, specifically the geostratum and a dichotomous variable defined to identify participants likely to have relatively high versus relatively low exposures (see section VIII.B.3.b.2). The analyses of disease and thyroid UDA outcomes in relation to these alternative exposure variables did not make use of the estimated doses. Nevertheless it was of interest to examine the distributions of estimated doses within the categories defined by these two variables.

The distributions of estimated doses are shown by geostratum in Table IX.B-4 above, and by the dichotomous exposure variable in Table IX.B-13 below. Note that the low exposure group included the 249 out-of-area participants for whom the CIDER program does not calculate a dose estimate. Therefore the description of the estimated dose distribution for the low exposure group in Table IX.B-13 refers only to the other 428 participants. As expected, estimated doses of participants in the low exposure group were generally lower, with a mean of 23 mGy, compared to the high exposure group with mean 288 mGy.

However there was substantial overlap in the distributions of estimated doses: the maximum estimated dose in the low exposure group was 160 mGy, while the minimum estimated dose in the high exposure group was 12 mGy. This overlap is not surprising, since the dichotomous exposure variable uses only part of the detailed full set of information that enters into CIDER's calculation of individual dose estimates.

Table IX.B-13. Summary of Estimated Radiation Doses (in mGy) to the Thyroid, by Dichotomous Exposure Variable

Exposure Group	No.	Median	Minimum	Maximum	Mean	St. Dev.
Low – in area	428	15	.003	160	23	25
Low – out of area	249	---	---	---	---	---
High	580	224	12	1235	288	214

The remaining 2183 living evaluable participants, who could not be classified into either the low or the high exposure group, and all of whom were among the in-area group, had estimated doses ranging from 0.003 mGy to 2823 mGy, with mean 173 mGy.

B.4 Implications for Statistical Power

The study's statistical power to detect an effect of ^{131}I from Hanford was determined primarily by the number of living evaluable participants and by the mean and variance of their doses. As described in

section V.A above, the final cohort definition was established in order to ensure a high likelihood that there would be a sufficient number of living evaluable participants and a dose distribution with a sufficiently large variance. Since the power to detect a dose-response of a given magnitude depends on the background rates, power was calculated for three exemplary outcomes corresponding to a range of background rates:

- Any Benign Thyroid Nodule, representing outcomes with intermediate background rates (assumed for power calculations to be 0.05 or 5% for women, 0.02 or 2% for men).

- Thyroid Carcinoma, representing outcomes with low background rates (assumed to be 0.007 or 0.7% for women, 0.003 or 0.3% for men).

- Ultrasound Detected Abnormalities, representing outcomes with high background rates (assumed to be 0.40 or 40% for both sexes).

Table IX.B-14 below summarizes the projections that were made for the Full Study cohort based on the results of the Pilot Study, assuming one-sided tests for a positive dose-response (i.e., slope > 0) at critical level $\alpha = 0.05$, and ignoring dose uncertainties. Also shown are the results that were actually obtained.

Table IX.B-14. Comparison of Projected and Obtained Statistical Power

	Projected	Obtained
Number of in-area living evaluable participants	3277	3191
Mean of dose distribution (mGy)	152	174
Variance of dose distribution (mGy^2)	38619	50150
Any benign thyroid nodule (intermediate background rates): power to detect 0.05 per Gy	0.91	0.95
Thyroid carcinoma (low background rates): power to detect 0.025 per Gy	0.93	0.96
UDAs (high background rates): power to detect 0.12 per Gy	0.86	0.92

It should be noted that the projected results assumed that out-of-area participants would be included, while the "Obtained" results are limited to the in-area participants, who were the basis for the primary analyses of the radiation dose-responses. Although the number of in-area living evaluable participants (3191) fell a bit short of the projection, the mean and variance of the dose distribution were larger than projected. As a result, the statistical power exceeded the projections.

As noted in the NRC's review of the draft HTDS Final Report, the uncertainties of the estimated doses could be expected to reduce the study's power from the levels summarized in Table IX.B-14 above (138). While it would be desirable to calculate the study's power with a direct adjustment for the dose uncertainties, this is impractical due to the complex nature of the correlations of the uncertainties between individual participants. Therefore, in order to assess the impact of dose uncertainty on the study's statistical power, a simulation analysis was performed. Such simulation studies are often used to investigate statistical power when exact calculations are impractical. The basic idea is to randomly generate ("simulate") a large number of data sets that mimic the key characteristics of the study (e.g., background rates, variance of the dose estimates, magnitudes and correlations of the dose uncertainties) for a specific hypothesis (null or alternative). Each simulated data set includes outcome data that are

themselves randomly generated under the hypothesis, and the significance of the resulting dose-response is tested. The proportion of data sets for which the null hypothesis is rejected is then an estimate of the study's power at the specific hypothesis.

For the HTDS, the simulation study began by specifying, for a given exemplary outcome (e.g., any benign thyroid nodule, thyroid carcinoma, or UDA as in Table IX.B-14 above), the sex-specific background rates and the slope of the dose-response under a specific hypothesis of interest (e.g., for any benign thyroid nodule, background rates 0.05 for women and 0.02 for men, and slope 0.05 per Gy). The simulation then proceeded through the following steps:

- Step 1: Randomly select 100 dose realizations with replacement from the existing 100 realizations. (Selection with replacement means that some realizations might not be selected, while others might be selected more than once.) Calculate each participant's median dose from the 100 randomly selected realizations.

- Step 2: Randomly select one of the 100 dose realizations. Treating this single set of doses as the "true" doses, calculate each participant's "true" probability of having the disease outcome using the sex-stratified linear model (see section VIII.C.1.a above) and the specified parameter values. Then randomly generate each participant's disease outcome (present or absent), with the probability of having the disease given by his or her "true" probability.
- Step 3: Fit the sex-stratified linear probability model using the median doses from Step 1 and the outcomes from Step 2, and determine whether the estimated slope is significantly greater than 0 at a given critical level (e.g., $\alpha = 0.05$ or 0.10).

After repeating Steps 1 through 3 for a large number of iterations (e.g., 1000), the proportion of iterations for which the estimated slope was significantly greater than 0 was calculated. This proportion is an estimate of the study's statistical power, i.e., of the probability of rejecting the null hypothesis that the slope is 0.

Note that the random selection with replacement in step 1 was used to ensure that this procedure accounted properly for not only the magnitude of the dose uncertainties, but for the between-participant correlations of dose uncertainties as well.

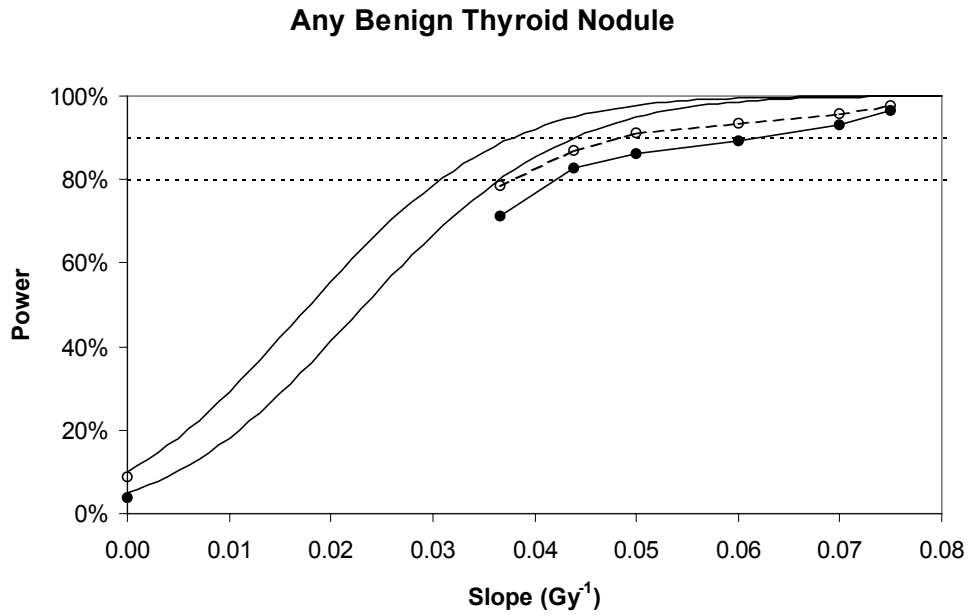
The results of the simulation study for the outcome of any benign thyroid nodule are shown in Table IX.B-15 and Figure IX.B-7. The power was evaluated for the following values of the slope parameter: 0 (i.e., the null hypothesis); 0.036 and 0.044 per Gy, for which the test at critical level $\alpha = 0.05$ has power 0.80 and 0.90, respectively, if dose uncertainty is ignored; and 0.05, 0.06, 0.07, and 0.75 per Gy. Background rates were assumed to be 0.05 for women and 0.02 for men. As expected, the dose uncertainties had no evident effect on the size of the test, i.e., the power under the null hypothesis (slope = 0). In addition, for alternative hypotheses with the slope greater than zero, the simulation study indicated that there was a modest loss of power due to dose uncertainties. For example, if the true slope of the linear dose-response is 0.05 per Gy (5% per Gy), then the estimated power of the test at critical level $\alpha = 0.05$ based on the simulation study (i.e., accounting for dose uncertainties) was 0.863, somewhat less than the value of 0.95 obtained if uncertainty was ignored.

Table IX.B-15. Effect of Dose Uncertainty on Statistical Power: Any Benign Thyroid Nodule

	Power of Test at Critical Level $\alpha = 0.05$		Power of Test at Critical Level $\alpha = 0.10$	
Slope (per Gy)	Ignoring Uncertainty	Accounting for Uncertainty	Ignoring Uncertainty	Accounting for Uncertainty
0	0.05	0.039	0.10	0.090
.036	0.80	0.713	0.88	0.787
.044	0.90	0.829	0.95	0.870
.050	0.95	0.863	0.98	0.912
.060	0.99	0.893	0.99	0.933
.070	0.997	0.931	0.999	0.956
.075	0.999	0.967	1.00	0.978

Although the study was designed to ensure that tests at critical level $\alpha = 0.05$ have adequate power, it should be recognized that dose-response parameters with p-values greater than 0.05 might also be considered evidence of a radiation effect. Therefore results are also shown in Table IX.B-15 for tests at critical level $\alpha = 0.10$. For example, the study had an estimated power of 0.912 for finding a dose-response with p-value < 0.10 if the true dose-response in fact had slope 0.05 per Gy.

Figure IX.B-7. Effect of Dose Uncertainty on Statistical Power: Any Benign Thyroid Nodule



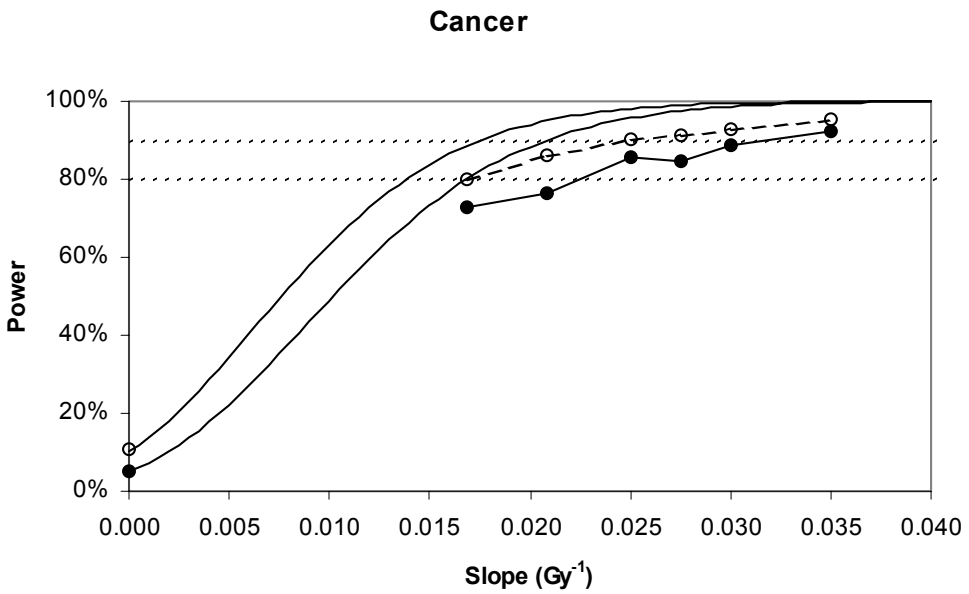
Solid lines show power calculated ignoring dose uncertainties for tests at critical level $\alpha = 0.05$ (lower curve) and 0.10 (upper curve). Circles show estimated power accounting for dose uncertainties, for tests at $\alpha = 0.05$ (solid) and 0.10 (open).

Table IX.B-16 and Figure IX.B-8 display similar results for the outcome of thyroid carcinoma, which represents outcomes with low background rates. The power was evaluated for the following values of the slope parameter: 0 (i.e., the null hypothesis); 0.0169 and 0.0208 per Gy, for which the test at critical level $\alpha = 0.05$ has power 0.80 and 0.90, respectively, if dose uncertainty is ignored; and 0.025, 0.0275, 0.030, and 0.035 per Gy. Background rates were assumed to be 0.007 for women and 0.003 for men. Accounting for dose uncertainties, the power of the test at critical level $\alpha = 0.05$ to detect an effect of 0.025 per Gy (2.5% per Gy) was estimated to be 0.855, compared to 0.96 if uncertainty was ignored.

Table IX.B-16. Effect of Dose Uncertainty on Statistical Power: Thyroid Carcinoma

	Power of Test at Critical Level $\alpha = 0.05$		Power of Test at Critical Level $\alpha = 0.10$	
Slope (per Gy)	Ignoring Uncertainty	Accounting for Uncertainty	Ignoring Uncertainty	Accounting for Uncertainty
0	0.05		0.10	
0.0169	0.80	0.726	0.89	0.799
0.0208	0.90	0.764	0.95	0.864
0.025	0.96	0.855	0.98	0.904
0.0275	0.98	0.848	0.99	0.911
0.030	0.99	0.888	0.995	0.926
0.035	0.996	0.922	0.999	0.952

Figure IX.B-8. Effect of Dose Uncertainty on Statistical Power: Thyroid Carcinoma



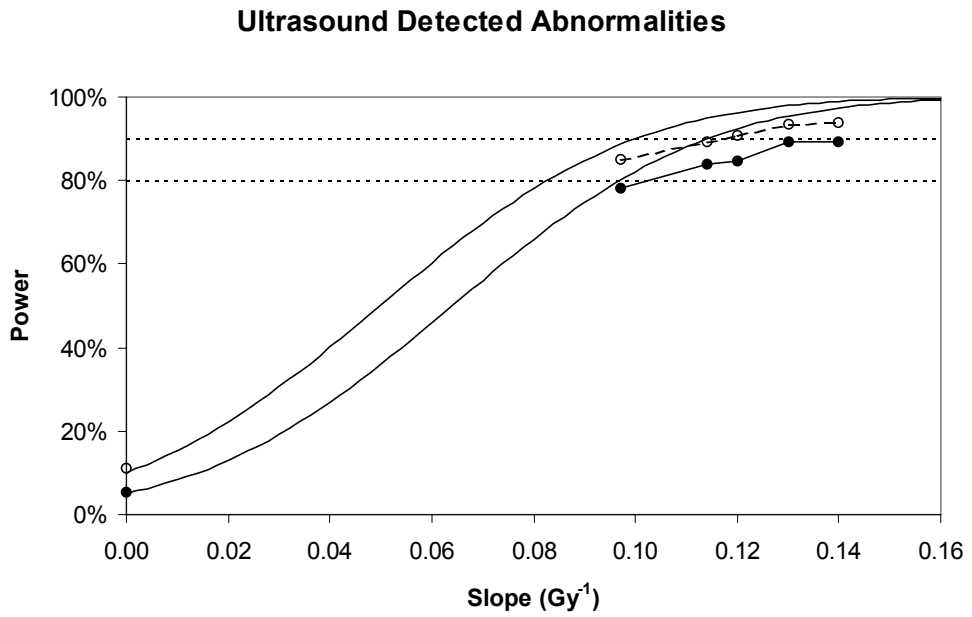
Solid lines show power calculated ignoring dose uncertainties for tests at critical level $\alpha = 0.05$ (lower curve) and 0.10 (upper curve). Circles show estimated power accounting for dose uncertainties, for tests at $\alpha = 0.05$ (solid) and 0.10 (open).

Table IX.B-17 and Figure IX.B-9 display similar results for the outcome of ultrasound detected abnormality, which represents outcomes with high background rates. The power was evaluated for the following values of the slope parameter: 0 (i.e., the null hypothesis); 0.097 and 0.114 per Gy, for which the test at critical level $\alpha = 0.05$ has power 0.80 and 0.90, respectively, if dose uncertainty is ignored; and 0.12, 0.13, and 0.14 per Gy. Background rates were assumed to be 0.40 for both women and men. Accounting for dose uncertainties, the power of the test at critical level $\alpha = 0.05$ to detect an effect of 0.12 per Gy (12% per Gy) was estimated to be 0.847, compared to 0.92 if uncertainty was ignored.

Table IX.B-17. Effect of Dose Uncertainty on Statistical Power: Ultrasound Detected Abnormalities

	Power of Test at Critical Level $\alpha = 0.05$		Power of Test at Critical Level $\alpha = 0.10$	
Slope (per Gy)	Ignoring Uncertainty	Accounting for Uncertainty	Ignoring Uncertainty	Accounting for Uncertainty
0	0.05	0.053	0.10	0.113
0.097	0.80	0.782	0.89	0.852
0.114	0.90	0.840	0.95	0.891
0.12	0.92	0.847	0.98	0.909
0.13	0.95	0.892	0.99	0.936
0.14	0.97	0.894	0.995	0.937

Figure IX.B-9. Effect of Dose Uncertainty on Statistical Power: Ultrasound Detected Abnormalities



Solid lines show power calculated ignoring dose uncertainties for tests at critical level $\alpha = 0.05$ (lower curve) and 0.10 (upper curve). Circles show estimated power accounting for dose uncertainties, for tests at $\alpha = 0.05$ (solid) and 0.10 (open).

In summary, the results of the simulation study showed that the effect of dose uncertainty was, as expected, to reduce the study's statistical power somewhat below the levels calculated with the uncertainties ignored. However, as summarized in Table IX.B-18 below, the reduction was modest, with about 85% power available for the alternative hypotheses to which the study's design was originally targeted.

Table IX.B-18. Summary of Effect of Dose Uncertainties on Statistical Power (one-sided tests at critical level $\alpha = 0.05$)

	Ignoring Uncertainty	Accounting for Uncertainty
Any benign thyroid nodule (intermediate background rates): power to detect 5% per Gy	0.95	0.863
Thyroid carcinoma (low background rates): power to detect 2.5% per Gy	0.96	0.855
UDAs (high background rates): power to detect 12% per Gy	0.92	0.847

To interpret the study's power properly, it is important to consider not only the level of power, but also the size of the dose-response effect for which that power is obtained. For example, for the exemplary outcome with low background rates, thyroid cancer, with assumed background probabilities of 0.7% and 0.3% for women and men, respectively, a linear dose-response with slope 2.5% per Gy yields probabilities of 1.1% and 0.7%, respectively, at the study participants' average dose of 174 mGy. These can also be expressed as relative risks of 1.1/0.7 = 1.62 and 2.45 for women and men, respectively, for an overall

average of 2.04. For the exemplary outcomes with intermediate (any benign thyroid nodule) or high (thyroid UDA) background rates, the corresponding relative risks (average over both sexes) are markedly smaller: 1.30 (5% per Gy) and 1.05 (12% per Gy), respectively. These represent the magnitudes of the effects for which the study's one-sided tests at critical level $\alpha = 0.05$ had estimates of about 85% to 86% power after accounting for the effects of dose uncertainties (see Table IX.B-18 above).

For comparison to results of other studies, the magnitudes of radiation effects can be expressed as the relative risks at 1000 mGy (1 Gy). For the low background rate example of thyroid cancer, a slope of 2.5% per Gy corresponds to probabilities of 3.2% and 2.8% for women and men at 1 Gy, respectively, i.e., to relative risks of 4.57 and 9.33, and an average of 6.95, at 1 Gy. This is similar to the estimated relative risk of 8.9 at 1 Gy reported for the Utah Study in their analysis that did not account for the effects of dose uncertainties (10139). However the appropriate comparison is to the estimated relative risk that is obtained after adjusting for the effect of dose uncertainties. The authors of the Utah Study reported that their uncertainty-adjusted estimates were about three-fold greater than the unadjusted estimates, corresponding to a relative risk of $1 + 3 \times (8.9 - 1)$, or about 25 at 1 Gy. A recent analysis suggested that the adjustment should perhaps be smaller: Mallick and colleagues analyzed the Utah Study's data concerning thyroid neoplasms and concluded that the estimated relative risk at 1 Gy should be approximately doubled, rather than tripled, to account for dose uncertainties (139). Assuming this conclusion applies to thyroid cancer, the estimated relative risk would be about 17 at 1 Gy. The HTDS clearly had adequate statistical power to detect an effect of this magnitude. For example, after accounting for dose uncertainty there was an estimated 92% power to detect a linear dose-response with a slope of 3.5% per Gy for thyroid cancer (Table IX.B-18 above), which corresponds to an average relative risk (both sexes combined) of 9.33 at 1 Gy, well below the estimated effect from the Utah Study.

B.5. Out-of-Area Participants

The numbers of out-of-area subjects are shown by sex, birth year, and geostratum in Table IX.B-19. The percentage of out-of-area participants was 7.2% for women (125/1747) and 7.3% for men (124/1693), but varied widely among birth years and geostrata.

Table IX.B-19. Proportions of Out-of-Area Participants, by Sex, Birth Year, and Geostratum

		Living Evaluable Participants	Out-of-Area	
			No.	%
Sex	Female	1747	125	7.2
	Male	1693	124	7.3
Birth Year	1940	243	25	10.3
	1941	283	28	9.9
	1942	472	37	7.8
	1943	560	76	13.6
	1944	906	77	8.5
	1945	611	3	0.5
	1946	365	3	0.8
Geostratum	Richland	352	4	1.1
	Pasco/Kennewick	1009	99	9.8
	Walla Walla City	264	14	5.3
	Benton Co.	734	78	10.6
	Franklin Co.	149	8	5.4
	Walla Walla Co.	334	14	4.2
	Okanogan Co.	139	14	10.1
	Ferry/Stevens Cos.	138	7	5.1
	Adams Co.	321	11	3.4
Total		3440	249	7.2

Only 6 (0.6%) of the 976 participants born in 1945 or 1946 were in the out-of-area group. In the earlier years, however, the percentage ranged from 7.8% (37/472) for 1942 to 13.6% (76/560) for 1943. The sharp drop in 1945-46 reflects that fact that the nearly all participants lived at or near their mother's "usual place of residence" for at least some time after their births. Consequently most participants born in 1945-46 first lived within the HEDR geographical domain. Participants born before 1945 and therefore, for the most part, before the start of ^{131}I releases from Hanford, had more time during which their families might move outside the HEDR domain.

Regarding geostrata, only 4 (1.1%) of the 352 participants in the Richland geostratum were in the out-of-area group. This occurred primarily because Richland was not defined as separate geostratum until 1944. Therefore it does not include participants born during 1940-1943 who, as explained above, had a greater likelihood of moving outside the HEDR domain before the start if Hanford's ^{131}I releases. In the other eight geostrata the percentage of out-of-area participants ranged from 3.4% (11/321) in the Adams County geostratum to 10.6% (78/734) in the Benton County geostratum.

C. Thyroid Cancer

C.1. Occurrence of Thyroid Cancer

The primary and alternative definitions for thyroid cancer were as follows:

- Primary definition: HTDS or prior histologic diagnosis (19 cases)
- Alternative definition: HTDS or prior histologic or clinical diagnosis (20 cases)

Twenty participants (0.6%) were diagnosed with thyroid cancer (Table IX.C-1), including 13 women (0.7%) and 7 men (0.4%). Of the twenty participants found to have thyroid cancer, all but one had diagnoses based on histologic evidence from either the HTDS examination (12) or prior medical care (7). Only one living evaluable participant's diagnosis of thyroid cancer was based on a prior clinical diagnosis. This participant's histology records had been destroyed, but her medical records from 1966 included mention of "Thyroidectomy (cancer) 4/65".

Of the 20 cancer diagnoses, 12 (60%) resulted from the HTDS examination and 8 (40%) were made prior to the participant's HTDS examination.

Table IX.C-1. Diagnoses of Thyroid Cancer, by Basis for Diagnosis and Sex

	Female		Male		Total	
Diagnosis of Thyroid Cancer	No.	%	No.	%	No.	%
Yes	13	0.7	7	0.4	20	0.6
▪ Histologic diagnosis: HTDS	6	0.3	6	0.4	12	0.3
▪ Prior histologic diagnosis	6	0.3	1	0.1	7	0.2
▪ Prior clinical diagnosis	1	0.1	0	--	1	0.0
No	1732	99.1	1685	99.5	3417	99.3
Unknown	2	0.1	1	0.1	3	0.1
Total	1747	100.0	1693	100.0	3440	100.0

Three additional living evaluable participants were classified "unknown" with regard to diagnosis of thyroid cancer. One of these participants had a fine needle aspiration (FNA) prior to the HTDS clinic of a mass outside of the thyroid. This mass was not seen or felt at the HTDS clinic, and no surgery was ever performed, thus thyroid cancer could not be ruled out. The second participant did not have a fine needle aspiration at the HTDS clinic due to a history of cardiac risk, and never had an FNA performed subsequent to the clinic. Again thyroid cancer could not be ruled out. For the third participant the two doctors at the HTDS clinic disagreed as to whether the subject had a lobulation or a small nodule and the ultrasound did not identify any nodules. These three participants were included as non-cases in analyses of the thyroid cancer dose-response.

Three other participants or potential participants had evidence of thyroid cancers that were not included in the primary analysis:

- Two living evaluable participants had thyroid cancers diagnosed after participating in HTDS. In one case the thyroid pathology was incidental to an HTDS recommendation for parathyroid surgery. In the other case the HTDS evaluation concluded that the two palpable nodules at the clinic were most likely non-thyroid, based on a normal nuclear scan. It was subsequently conveyed via a phone call from the participant that she had thyroid cancer. It was determined that since our evaluation of this subject was concluded with no recommendation for surgery or follow-up for definitive pathology, the information from the phone call could not be used or pursued. Although these two diagnoses could not be used in the primary analysis of thyroid cancer, they were included in an additional dose-response analysis (see section IX.C.2.c below).

- One potential participant who refused to participate in HTDS gave as a reason that he/she had thyroid cancer and had already seen too many doctors. Since this person did not participate in the HTDS, this case could not be included in any analyses.

Sixteen (80.0%) of the 20 cancer cases had papillary cancer, while three (15.0%) had follicular cancer (Table IX.C-2). The histologic type was unknown for the one participant with only a prior clinical diagnosis.

Table IX.C-2. Frequency Distribution of Histologic Types of Thyroid Cancer, by Sex

	Female		Male		Total	
Histologic Type	Cases	%	Cases	%	Cases	%
Papillary Cancer	10	76.9	6	85.7	16	80.0
Follicular Cancer	2	15.4	1	14.3	3	15.0
Unknown	1	7.7	0	--	1	5.0
Total	13	100.0	7	100.0	20	100.0

C.1.a Pathways to Diagnosis of Thyroid Cancer

The section above described the sources of information for all diagnoses of thyroid cancer among the living evaluable study participants. The diagnoses that resulted from the HTDS clinical examinations can also be characterized according to the method of detection (or "pathway to diagnosis"). As described in section V.F above, the HTDS employed a comprehensive diagnostic design in which participants received a thyroid ultrasound scan that was viewed only after two independent thyroid physical examinations were conducted by thyroid specialists. Additional thyroid exams were then conducted only if the ultrasound showed abnormalities that were not detected by the physicians. For the 12 diagnoses of thyroid cancer that were made as a result of the HTDS examination, Table IX.C-3 shows which component of the diagnostic process was instrumental in making the diagnosis. The majority of the thyroid cancers (10 or 83%) were detected because one or both of the physicians palpated a new thyroid mass before viewing the videotaped recording of the ultrasound examination. However the other two thyroid cancers were detected only when the physicians repeated the physical examination after reviewing the ultrasound scan. These descriptive results illustrate the contributions of multiple diagnostic methods in the evaluation process. They also underscore the differences that can occur in the prevalence of thyroid disease from one study to another depending on the diagnostic methods used.

Table IX.C–3. Pathways to Diagnosis of Thyroid Cancer

	Thyroid Cancer	
Pathway to Diagnosis	No.	%
Palpable prior to ultrasound	10	83.3
Palpable only after ultrasound	2	16.7
Palpable only (not detected on ultrasound)	0	--
Nonpalpable (detected only on ultrasound)	0	--
Total	12	100

C.2. Analysis of Thyroid Cancer Risk

C.2.a. Primary Analysis

Nineteen living evaluable participants had diagnoses of thyroid cancer based on HTDS or prior histologic evidence. Five of these cases were out-of-area participants, for whom the CIDER program could not calculate dose estimates. The numbers of cases and proportions with thyroid cancer are shown by sex and dose category in Table IX.C-4.

Table IX.C-4. Diagnoses of Thyroid Cancer by Sex, Dose Category, and Basis for Diagnosis

A. Female

Thyroid Radiation Dose	Living Evaluable Female	Primary Definition: Cases Based on HTDS or Prior Histologic Diagnosis		Alternative Definition: Cases Based on HTDS or Prior Histologic or Clinical Diagnosis	
(mGy)	No.	No.	%	No.	%
Out of Area	125	2	1.6	3	2.4
< 10	182	1	0.5	1	0.5
10-49	320	3	0.9	3	0.9
50-99	313	1	0.3	1	0.3
100-149	220	1	0.5	1	0.5
150-199	126	1	0.8	1	0.8
200-299	139	1	0.7	1	0.7
300-399	144	1	0.7	1	0.7
400-999	171	1	0.6	1	0.6
1000+	7	0	--	0	--
Total	1747	12	0.7	13	0.7

B. Male

Thyroid Radiation Dose	Living Evaluable Male	Primary Definition: Cases Based on HTDS or Prior Histologic Diagnosis		Alternative Definition: Cases Based on HTDS or Prior Histologic or Clinical Diagnosis	
(mGy)	No.	No.	%	No.	%
Out of Area	124	3	2.4	3	2.4
< 10	186	1	0.5	1	0.5
10-49	314	2	0.6	2	0.6
50-99	310	0	--	0	--
100-149	171	0	--	0	--
150-199	109	0	--	0	--
200-299	148	0	--	0	--
300-399	160	0	--	0	--
400-999	154	0	--	0	--
1000+	17	1	5.9	1	5.9
Total	1693	7	0.4	7	0.4

The highest estimated dose among the 14 in-area cases was 1083 mGy. Parameter estimates for the linear dose-response model based on the 3191 living evaluable in-area participants are shown in Table IX.C-5 below. Based on maximum likelihood analysis of the sex-stratified linear probability model, and using primary dose estimates (Table IX.C-5, row 1), the risk of thyroid cancer did not increase significantly with estimated dose ($p = 0.25$), with an estimated slope B of 0.002 per Gy, and 95% CI ranging from less than −0.001 to 0.017 per Gy. The background thyroid cancer rates were estimated to be 0.006 with confidence interval (0.001, 0.011) for women, and 0.002 with confidence interval (0, 0.005) for

men. Results obtained by least squares analysis using ungrouped or grouped data were similar (rows 2 and 3 of Table IX.C-5).

Table IX.C-5. Dose-Response Results for Diagnoses of Thyroid Cancer Based on HTDS or Prior Histologic Evidence

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
1.	Primary definition (HTDS or prior, histologic diagnosis)	Linear	Primary	None	MLE	.006 ± .002 (.001, .011)	.002 ± .001 (0*, .005)	.002 ± .004 (< –.001, .017)	0.25
2.	Primary definition	Linear	Primary	None	LSU	.005 ± .002 (.001, .010)	.002 ± .002 (0*, .006)	.005 ± .005 (–.008, .017)	0.19
3.	Primary definition	Linear	Primary	None	LSG	.006 ± .002 (.002, .011)	.003 ± .002 (0*, .007)	–.000 ± .006 (–.015, .014)	0.51
4.	Primary definition	Linear	Primary	+ 2 Incidental cases	MLE	.007 ± .002 (.002, .012)	.002 ± .001 (0*, .005)	.002 ± .004 (< –.001, .017)	0.28
5.	Primary definition	LQ	Primary	None	LSU	.006 ± .002 (.001, .011)	.002 ± .002 (0*, .007)	Lin: .002 ± .009 (–.020, .024) Quad: .002 ± .006 (–.012 .017)	Quad: 0.70

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model. "0*" indicates that the lower confidence limit for a background rate was less than 0.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.C-5. Dose-Response Results for Diagnoses of Thyroid Cancer Based on HTDS or Prior Histologic Evidence (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
6.	Primary definition	Logistic	Primary	None	MLE	.005 (.002, .013)	.002 (.001, .008)	.71 ± .79 (-1.18, 2.61)	0.22
7.	Primary definition	Linear	Primary	Exclude dose > 1000 mGy	MLE	.007 ± .002 (.001, .013)	.002 ± .001 (0*, .005)	−.002 ± .006 (NE, >.011)	0.77
8.	Primary definition	Logistic	Primary	Exclude dose > 400 mGy	MLE	.007 ± .003 (0*, .014)	.002 ± .001 (0*, .005)	−.006 ± .016 (NE, .015)	0.87
9.	Primary definition	Linear	Primary	Exclude OK and F/S geostrata	MLE	.006 ± .002 (.0005, .011)	.002 ± .001 (0*, .006)	.002 ± .004 (<-.001,.018)	0.26
10.	Primary definition	Linear	Alt. #1	None	MLE	.006 ± .002 (.001, .012)	.003 ± .002 (0*, .007)	−.001 ± .005 (NE, .015)	0.59

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model. "0*" indicates that the lower confidence limit for a background rate was less than 0.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.C-5. Dose-Response Results for Diagnoses of Thyroid Cancer Based on HTDS or Prior Histologic Evidence (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
11.	Primary definition	Linear	Alt. #2	None	MLE	006 ± .002 (.0005, .012)	.003 ± .001 (0*, .006)	–.001 ± .010 (NE, .008)	0.80
12.	Primary definition	Linear	Primary	Include OOA (scoping analysis #1)	MLE	.007 ± .002 (.002, .012)	.004 ± .002 (.0001, .008)	.0006 ± .004 (<-.002, >.015)	0.44
13.	Primary definition	Linear	Primary	Include OOA (scoping analysis #2)	MLE	.007 ± .002 (.002, .012)	.004 ± .002 (.0001, .008)	.0002 ± .004 (<-.002, >.014)	0.48

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model. "0*" indicates that the lower confidence limit for a background rate was less than 0.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

C.2.b. Alternative Definition for Diagnosis of Thyroid Cancer

As described in section IX.C.1 above, only one participant had a diagnosis of thyroid cancer based on anything other than HTDS or prior histologic diagnosis. This case was an out-of-area participant, and therefore had no effect on the primary dose-response analysis.

C.2.c. Effect of Including Incidental Diagnoses of Thyroid Cancer

As described in section IX.C.1 above, two living evaluable participants had thyroid cancers that were determined to be incidental. That is, each diagnosis was made after the participant's HTDS examination, and not as a result of a study recommendation for further evaluation of a possible thyroid cancer. These two cases were not included in the primary analysis of thyroid cancer to avoid introducing a possible reporting bias. However, in view of the importance of thyroid cancer as a disease outcome, additional analyses that included these two incidental cases were performed. These two participants were both in the in-area group, and their estimated thyroid radiation doses were 169 and 62 mGy. When these two incidental cases were included along with the 14 in-area cases in the primary analysis, the results were essentially unchanged, with estimated slope 0.002 per Gy with 95% CI ranging from less than –0.001 to 0.017 per Gy (Table IX.C-5, row 4).

C.2.d. Alternative Dose-Response Functions

Shown in row 5 of Table IX.C-5, the estimated regression coefficient for the dose-squared term in the linear-quadratic dose-response model [5] was 0.002 with Bonferroni-adjusted 95% confidence interval ranging from –0.012 to 0.017. Thus the addition of a quadratic term did not significantly improve the fit of the model ($p = 0.70$).

The regression parameter for the effect of dose in the sex-stratified logistic regression model [2] was estimated as 0.71 with Bonferroni-adjusted 95% confidence interval ranging from –1.18 to 2.61. Thus there was no evidence from the logistic regression model that cumulative incidence of thyroid cancer increased significantly with increasing dose ($p = 0.22$, Table IX.C-5, row 6).

C.2.e. Effect of Excluding Participants in High Dose Categories

The proportions of in-area women with cancer varied little over the dose categories shown in Table IX.C-4, ranging between 0.3% and 0.9%, with no cases among the seven women with doses over 1000 mGy. One of the four male thyroid cancer cases in the in-area group had an estimated dose of 1083 mGy, while the other three had doses less than 50 mGy. Consequently, when participants in the highest dose categories (> 1000 mGy or > 400 mGy) were excluded, the estimated slope of the dose-response decreased slightly, to –0.002 per Gy with Bonferroni-adjusted 95% upper confidence limit exceeding 0.011 per Gy among those with doses < 1000 mGy, and to –0.006 per Gy with upper confidence limit 0.015 per Gy among those with doses < 400 mGy (Table IX.C-5, rows 7 and 8). Thus there was no evidence that the dose-response results were inordinately influenced by the outcomes of participants in the highest dose categories.

C.2.f. Effect of Excluding Okanogan and Ferry/Stevens Geostrata

When participants in the Okanogan and Ferry/Stevens geostrata were excluded, the estimated slope of the dose-response changed only slightly, to 0.002 per Gy with 95% CI ranging from less than

–0.001 to 0.018 per Gy (Table IX.C-5, row 9). Thus there was no evidence that the dose-response results were inordinately influenced by the outcomes of participants in these geostrata.

C.2.g. Analysis of Thyroid Cancer in Relation to Alternative Dose Estimates

Parameter estimates for the linear dose-response model using the alternative dose estimates are shown in rows 10 and 11 of Table IX.C-5 above. For both alternative dose estimates the estimated slope B decreased as compared to the primary dose set, from 0.002 to –0.001, and thus in neither case was there evidence that the cumulative incidence of thyroid cancer increased with increasing dose.

C.2.h. Scoping Analyses Regarding Out-of-Area Participants

See section VIII.C.1.a.3 for a description of the scoping analyses that were performed to assess the possible impact of including the 249 out-of-area participants. As summarized in rows 12 and 13 of Table IX.C-5, in both analyses the inclusion of the out-of-area participants slightly decreased the estimated slope of the dose-response, but did not materially change the dose-response results.

C.2.i. Analysis of Thyroid Cancer in Relation to Alternative Representations of Exposure

In the analyses by geostratum and by dichotomous exposure variable, the sex and age-adjusted comparisons of cumulative incidence were performed as described in section VIII.C.2.a.2.

C.2.i.1. Analysis by Geostratum

Since only 19 participants had thyroid cancer (see Table IX.C-6), the test for heterogeneity among the nine geostrata had little statistical power. Therefore the absence of significant heterogeneity (p = 0.73) was not strong evidence against the possibility that the cumulative incidence of thyroid cancer might in fact vary among the geostrata. The percentages with cancer were somewhat higher in the Okanogan and Ferry/Stevens geostrata (1.4% for women, 0.7% for men) than in the remaining geostrata (0.6% and 0.4%), but this difference was also not statistically significant (p = 0.26).

Table IX.C-6. Diagnoses of Thyroid Cancer Based on HTDS or Prior Histologic Evidence, by Geostratum and Sex

	Female			Male			Total		
Geostratum	No.	Cases	%	No.	Cases	%	No.	Cases	%
Richland	179	1	0.6	173	1	0.6	352	2	0.6
Pasco/Kennewick	508	3	0.6	501	1	0.2	1009	4	0.4
Benton County	376	2	0.5	358	2	0.6	734	4	0.5
Franklin County	73	0	--	76	0	--	149	0	--
Adams County	165	1	0.6	156	0	--	321	1	0.3
Walla Walla (city)	133	1	0.8	131	1	0.8	264	2	0.8
Walla Walla County	170	2	1.2	164	1	0.6	334	3	0.9
Okanogan County	75	1	1.3	64	1	1.6	139	2	1.4
Ferry/Stevens Counties	68	1	1.5	70	0	0.0	138	1	0.7
Total	1747	12	0.7	1693	7	0.4	3440	19	0.6

C.2.i.2. Analysis by Dichotomous Exposure Variable

See section VIII.B.3.b.2 above for a description of the high and low exposure categories. Eleven (0.9%) of the 1257 participants included in these analyses had thyroid cancer based on an HTDS or prior histologic examination (see Table IX.C-7). These included 3/580 (0.5%) in the high exposure group and 8/677 (1.2%) in the low exposure group. Thus there was no evidence that cumulative incidence of thyroid cancer was elevated in the high exposure group (p = 0.86).

Table IX.C-7. Diagnoses of Thyroid Cancer based on HTDS or prior histologic evidence, by exposure group and sex

Exposure	Female			Male			Total		
Group	No.	Cases	%	No.	Cases	%	No.	Cases	%
Low	351	4	1.1	326	4	1.2	677	8	1.2
High	298	2	0.7	282	1	0.4	580	3	0.5
Total	649	6	0.9	608	5	0.8	1257	11	0.9

C.2.j. Confounding and Effect Modification

There were too few participants with diagnoses of thyroid cancer to warrant any analysis of confounding or effect modification.

C.2.k. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for thyroid cancer are shown in Figure IX.C-1 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in that figure are calculated at the 98.33% confidence interval, i.e., are adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates. While the point estimate of the slope is greater than 0 for 65 of the 100 realizations, the confidence interval includes 0 for all of the 100 realizations. Also shown in Figure IX.C-1 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean and mean of each participant's 100 dose realizations. In summary, for none of the 100 realizations of the estimated doses was there a statistically significant dose-response.

Figure IX.C-1. Plot of Estimated Slope by Dose Realization

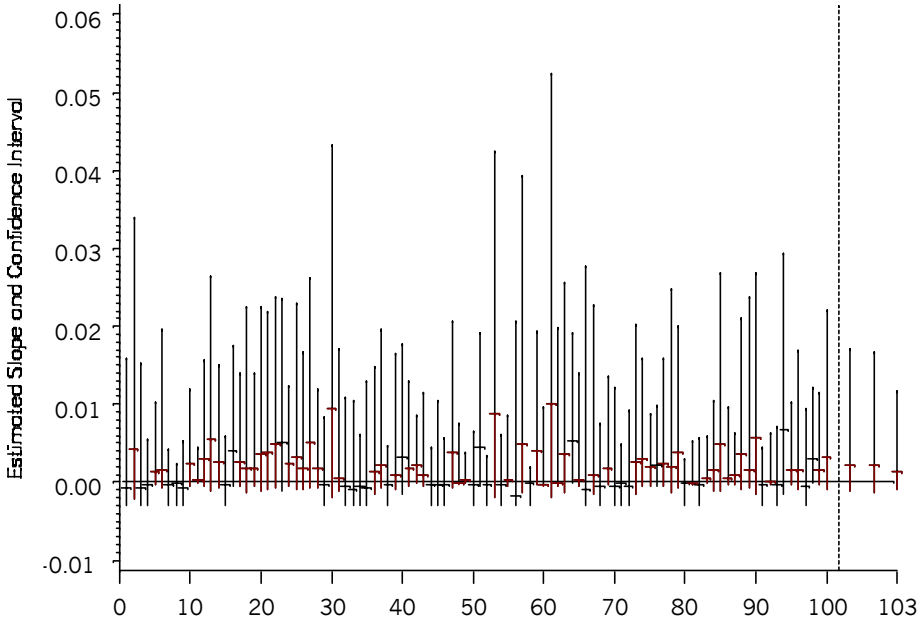
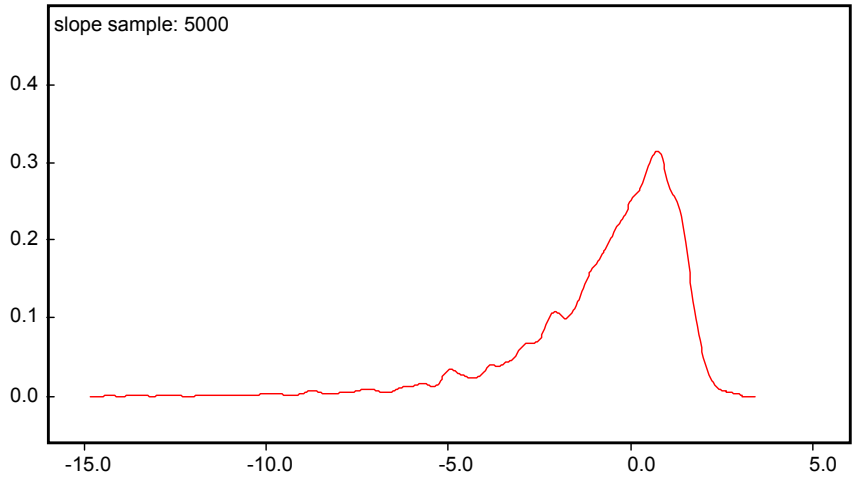


Figure IX.C-2 displays the distribution of the 5000 estimates of the logistic coefficient obtained by the simulation procedures described in section VIII.C.2.b.3 above. It is evident from the figure that most of the estimates were between about –5.0 and 2.0. The estimate was less than or equal to 0 for 2574 of the 5000 replications, implying an empirical one-tailed p-value of 0.52. The median estimate was –0.06, and the upper and lower percentiles corresponding to the Bonferroni-adjusted 95% confidence interval were –8.29 and 2.11. These may be compared to the estimate of 0.71 with confidence interval (-1.18, 2.61) obtained using the median dose estimates without adjustment for uncertainty. Thus, this method of adjusting the estimated logistic regression coefficient for the uncertainty in the dose estimates did not provide evidence that the cumulative incidence of thyroid cancer increased with increasing dose.

Figure IX.C-2. Distribution of Simulation Estimates of Logistic Regression Coefficient



D. Benign Thyroid Nodule

D.1. Occurrence of Benign Thyroid Nodule

The primary and alternative definitions for benign thyroid nodule were as follows:

- Primary definition: HTDS or prior, histologic or cytologic diagnosis (249 cases)
- Alternative definition #1: HTDS or prior, histologic, cytologic or clinical diagnosis (287 cases)
- Alternative definition #2: Any diagnosis or participant/respondent report (297 cases).

Table IX.D-1 shows the numbers and percentages of living evaluable participants with diagnoses of benign thyroid nodule, and the bases for those diagnoses, by sex. Two hundred and forty-nine (7.2%) living evaluable participants had a diagnosis of benign thyroid nodule based on histologic or cytologic evidence arising from the HTDS examination or from a prior diagnosis, with 170 (9.7%) women 79 (4.7%) men having this condition, respectively. Thirty-eight (1.1%) participants had diagnoses classified as clinical. Additionally, for 10 (0.3%) the diagnosis was based solely on a report by the participant or his/her CATI respondent.

Table IX.D-1. Diagnoses of Benign Thyroid Nodule, by Basis for Diagnosis and Sex

	Female		Male		Total	
Diagnosis of Benign Thyroid Nodule	No.	%	No.	%	No.	%
Yes	200	11.4	97	5.7	297	8.6
▪ Histologic diagnosis: HTDS	7	0.4	6	0.4	13	0.4
▪ Cytologic diagnosis: HTDS	142	8.1	65	3.8	207	6.0
▪ Prior histologic diagnosis	19	1.1	7	0.4	26	0.8
▪ Prior cytologic diagnosis	2	0.1	1	0.1	3	0.1
▪ Clinical diagnosis: HTDS	16	0.9	13	0.8	29	0.8
▪ Prior clinical diagnosis	7	0.4	2	0.1	9	0.3
▪ Participant/respondent report	7	0.4	3	0.2	10	0.3
No	1545	88.4	1595	94.2	3140	91.3
Unknown	2	0.1	1	0.1	3	0.1
Total	1747	100.0	1693	100.0	3440	100.0

Three living evaluable participants were classified "unknown" with regard to diagnosis of benign thyroid nodule. One of these participants reported a history of having a thyroid "lump" removed, but had no record of surgery or evidence of a surgical scar. The second participant's medical record included mention of thyroid nodule by one physician. However a second physician disagreed, revising the diagnosis to thyromegaly, with decrease in size after treatment. For the third participant, the two doctors at the HTDS clinic disagreed as to whether the subject had a lobulation or a small nodule and the ultrasound did not identify any nodules (this participant was also classified as "unknown" for diagnosis of thyroid cancer). These three participants were included as non-cases in analyses of the dose-response for benign thyroid nodule.

As shown in Table IX.D-2, the majority of benign thyroid nodules were colloid nodules (69.7%). Follicular adenomas accounted for only 4.7% of the diagnoses. The remaining 33.0% included a variety of types of nodules, which are described in Table IX.D-3.

Table IX.D-2. Frequency Distribution of Histologic/Cytologic Types of Benign Thyroid Nodule, by Sex

	Female		Male		Total	
Histologic/Cytologic Type	Cases	%	Cases	%	Cases	%
Colloid nodule	139	69.5	68	70.1	207	69.7
Follicular adenoma	8	4.0	6	6.2	14	4.7
Other	71	35.5	26	26.8	97	32.7
Total with benign thyroid nodule	200	100.0	97	100.0	297	100.0

Note: A participant can have >1 histologic/cytologic type

Table IX.D-3. Frequency Distribution of Other Histologic/Cytologic Types of Benign Thyroid Nodule, by Sex

	Female		Male		Total	
Other Histologic/Cytologic Type	Cases	%	Cases	%	Cases	%
Unknown/uncertain*	25	35.2	14	53.8	39	40.2
Hashimoto's thyroiditis	23	32.4	6	23.1	29	29.9
Thyroglossal duct cyst	4	5.6	1	3.8	5	5.2
Adenomatous nodule/goiter	3	4.2	2	7.7	5	5.2
Benign follicular nodule	7	9.9	1	3.8	8	8.2
Benign nodular goiter	2	2.8	0	--	2	2.1
Chronic thyroiditis w/benign follicles & Hurthle cells	1	1.4	0	--	1	1.0
Colloid nodule vs follicular adenoma	1	1.4	0	--	1	1.0
Hashimoto's & non-neoplastic follicular nodule w/colloid	1	1.4	0	--	1	1.0
Nodular hyperplasia	1	1.4	0	--	1	1.0
Possible thyroiditis	1	1.4	0	--	1	1.0
Simple cyst	1	1.4	0	--	1	1.0
Nondiagnostic, probable colloid nodule	1	1.4	0	--	1	1.0
Unknown due to participant/respondent report	0	--	1	3.8	1	1.0
Probable neoplastic macrofollicular nodule	0	--	1	3.8	1	1.0
Total with other histologic/cytologic type	71	100.0	26	100.0	97	100.0

* No cytology available

Of the 98 participants with histologic/cytologic type classified 'Other' (Table IX.D-3), 39 (39.8%) were of unknown or uncertain type, meaning no cytology was available. Another 29 (29.6%) were associated with Hashimoto's thyroiditis, 8 (8.2%) were due to a benign follicular nodule, 5 (5.1%) were due to a thyroglossal duct cyst, 5 (5.1%) were due to an adenomatous nodule, 2 (2.0%) were due to a benign nodular goiter, and the remaining 9 were due to varying individual specifications of the histologic/cytologic type.

D.1.a. Additional Disease Outcomes Related to Benign Thyroid Nodule

The following additional disease outcomes related to benign thyroid nodule were considered. These outcomes were defined based on the primary definition of benign thyroid nodule (i.e. HTDS or prior, histologic or cytologic evidence).

D.1.a.1 Benign Thyroid Nodules and Nodules Suspicious for Thyroid Follicular Adenoma

Additional analyses were performed in which the participants with either benign thyroid nodules or nodules coded as "suspicious for follicular neoplasm" were combined as cases. The category of suspicious for follicular neoplasm deserves some additional comment. Participants having FNA biopsy for a palpable nodule or a nonpalpable nodule larger than an average of 1.5 cm, were recommended to have further evaluation or consideration of thyroid surgery if the FNA result was reported as either suspicious for malignancy or suspicious for follicular neoplasm. For those participants who did have surgery, the HTDS final diagnosis was then designated as either cancer or benign thyroid nodule based on the surgical pathology. However, there were 16 participants with FNA results reported as suspicious for follicular neoplasm who chose not to have surgery. None of those individuals had FNA results that were suspicious for cancer. Their FNA results showed either intermediate or high probability of follicular neoplasm; none were suspicious for papillary cancer. Although these 16 cases were most likely to represent a benign thyroid nodule, the risk of thyroid cancer in such cases has been reported to be approximately 10-30%.

Sixteen participants without other benign thyroid nodules (14 women, 2 men) had diagnoses of nodules suspicious for follicular neoplasm, all based on cytology. Consequently the 3440 living evaluable participants included 265 (7.7%) with diagnoses of benign thyroid nodule or nodule suspicious for follicular neoplasm (Table IX.D-4), with more than twice as many cases among women (10.5%) than men (4.8%).

Table IX.D-4. Benign Thyroid Nodule and Nodules Suspicious for Follicular Neoplasm, by Sex

Benign Thyroid Nodule or Nodule Suspicious for Follicular Neoplasm	Female		Male		Total	
	No.	%	No.	%	No.	%
Yes	184	10.5	81	4.8	265	7.7
No	1563	89.5	1612	95.2	3175	92.3
Total	1747	100.0	1693	100.0	3440	100.0

D.1.a.2. Benign Thyroid Nodule Excluding Non-neoplastic Disease

The outcome of benign thyroid nodule excluding non-neoplastic etiology was defined in order to exclude cases that might have a specific non-neoplastic etiology, as their inclusion might mask a dose-response effect. This outcome was defined to include participants with a diagnosis of benign thyroid nodule based on histologic or cytologic evidence from the HTDS or prior examination, but excluding those with any of the following:

- Autoimmune thyroiditis based on HTDS evaluation or medical records with supporting documentation;
- Graves disease based on HTDS evaluation or medical records with supporting documentation; or
- Hyperthyroidism based on HTDS evaluation or medical records with supporting documentation with an etiology of toxic nodular goiter or solitary toxic nodule.

Among the 3440 living evaluable participants 175 (5.1%) had a diagnosis of benign thyroid nodule excluding a non-neoplastic etiology, with the percentage of cases about twice as high for women (6.7%) as for men (3.4%) (Table IX.D-5).

Table IX.D-5. Benign Thyroid Nodule Excluding Non-neoplastic Disease, by Sex

Benign Thyroid Nodule Excluding Non-neoplastic Disease	Female		Male		Total	
	No.	%	No.	%	No.	%
Yes	117	6.7	58	3.4	175	5.1
No	1630	93.3	1635	96.6	3265	94.9
Total	1747	100.0	1693	100.0	3440	100.0

D.1.a.3. Solitary Benign Thyroid Nodule Detected without Ultrasound

The outcome of palpable, solitary, benign thyroid nodule detected without ultrasound was defined in order to simulate the effect of screening for thyroid disease by palpation only, i.e., without ultrasound examination. A total of 88 living evaluable participants (64 women, 24 men) had diagnoses of such nodules (Table IX.D-6).

Table IX.D-6. Solitary Benign Thyroid Nodule Detected without Ultrasound, by Sex

Solitary Benign Thyroid Nodule Detected without Ultrasound	Female		Male		Total	
	No.	%	No.	%	No.	%
Yes	64	3.7	24	1.4	88	2.6
No	1683	96.3	1669	98.6	3352	97.4
Total	1747	100.0	1693	100.0	3440	100.0

For the majority of the 88 living evaluable participants with solitary benign thyroid nodules that were detected without ultrasound, i.e., by palpation, those nodules were also observed on the ultrasound examination. However for 21 (24%) of the 88, those nodules were not detected by ultrasound. Twelve (57%) of these 21 participants each had 1-6 discrete focal ultrasound abnormalities in addition to the palpable nodule which was not detected on ultrasound. In addition, 15 of 21 (71%) had documented Hashimoto's thyroiditis. Only 4 participants (0.1% of the 3429 living evaluable participants whose thyroid glands were visible in their ultrasound examinations) had a palpable nodule with a completely normal ultrasound scan. These results suggest that the reason for the discordance between palpation and ultrasound in this small group was the abnormal thyroid tissue that is present throughout the gland in individuals with Hashimoto's thyroiditis, a fact well known in clinical practice. Since only 4 participants had true palpable nodules that were not detected by ultrasound, a dose-response analysis of this specific outcome was not feasible.

D.1.a.4. Benign Thyroid Nodule Excluding Colloid-Only Nodules

In the primary analysis, thyroid nodules with abundant colloid but insufficient follicular cells (designated for this study as "colloid-only" nodules) were classified as benign thyroid nodules. Since such a cytology result is technically nondiagnostic, an additional analysis was performed in which the colloid-only nodules were not counted among the benign thyroid nodules. Of the 249 living evaluable participants with diagnoses of benign thyroid nodules, 18 (12 women and six men) had diagnoses based solely on colloid-only nodules. Thus a total of 231 (6.7%) had benign thyroid nodules excluding colloid-only nodules (Table IX.D-7).

Table IX.D-7. Benign Thyroid Nodule Excluding Colloid-Only Nodules, by Sex

Benign Thyroid Nodule Excluding Colloid-only Nodules	Female		Male		Total	
	No.	%	No.	%	No.	%
Yes	158	9.0	73	4.3	231	6.7
No	1589	91.0	1620	95.7	3209	93.3
Total	1747	100.0	1693	100.0	3440	100.0

D.1.a.5. Benign Colloid Nodules

Colloid nodules comprised the largest category of benign thyroid nodules. Thus the outcome of benign colloid nodules was defined to determine whether colloid nodules might be related to ^{131}I exposure. Participants were counted as cases for this outcome if they had colloid nodules, regardless of whether they had any other benign thyroid nodules. As shown in Table IX.D-8 below, 201(5.8%) of the 3440 living evaluable participants had benign colloid nodules.

Table IX.D-8. Benign Colloid Nodules, by Sex

Benign Colloid Nodules	Female		Male		Total	
	No.	%	No.	%	No.	%
Yes	136	7.8	65	3.8	201	5.8
No	1611	92.2	1628	96.2	3239	94.2
Total	1747	100.0	1693	100.0	3440	100.0

D.1.b. Pathways to Diagnosis of Benign Thyroid Nodules and Thyroid Nodules Suspicious for Follicular Neoplasm

The diagnoses described above were based primarily on diagnostic testing done at the HTDS clinics as well as the participants' prior medical records. As was done for thyroid cancer, the diagnoses that resulted from the HTDS clinical examinations were characterized according to the method of detection (or "pathway to diagnosis"). As described in section V.F above, the HTDS employed a comprehensive diagnostic design in which participants received a thyroid ultrasound scan that was viewed only after two independent thyroid physical examinations were conducted by thyroid specialists. Additional thyroid examinations were then conducted only if the ultrasound showed abnormalities that were not detected by the physicians.

Table IX.D-9. shows the method of detection for diagnoses of benign thyroid nodules, or nodules suspicious for follicular neoplasm, that resulted from HTDS examinations.

Table IX.D-9. Pathways to Diagnosis of Benign Thyroid Nodules and Thyroid Nodules Suspicious for Follicular Neoplasm

	Benign Thyroid Nodule		Suspicious for Follicular Neoplasm		Total	
Pathway To Diagnosis	No.	%	No.	%	No.	%
Palpable prior to ultrasound	104	47.3	7	41.2	110*	46.6
Palpable only after ultrasound	67	30.5	7	41.2	74	31.4
Palpable only (not detected on ultrasound)	15	6.8	0	--	15	6.4
Nonpalpable (detected only on ultrasound)	28	12.7	2	11.8	30	12.7
Uncertain consensus on physician exam	0	--	1	5.9	1	0.4
Complex cases: FNA decision based on combination of ultrasound and palpation	6	2.7	0	--	6	2.5
Total	220	100	17	100	236	100

* Note that one participant with both a benign thyroid nodule and a nodule suspicious for follicular neoplasm, both of which were palpable prior to ultrasound, is only counted once in the Total column.

The results in Table IX.D-9 show that about half of these diagnoses (125 or 51%) could have been detected by palpation alone. However nearly a third of these diagnoses (74 or 31%) required ultrasound review before they were detected by palpation. For 30 (13%) of these diagnoses, ultrasound was the only method that led to the diagnosis; these cases were relatively large, nonpalpable nodules (>1.5 cm in 3 dimensions) that were biopsied because of their size. None of these cases showed thyroid cancer. The relative frequencies of the various pathways to diagnosis were about the same for nodules suspicious for follicular neoplasm as for diagnoses of benign thyroid nodules. As indicated previously for thyroid cancer, these descriptive results illustrate the contributions of multiple diagnostic methods in the evaluation process. They also underscore the fairly large differences that can occur in the prevalence of thyroid disease from one study to another depending on the diagnostic methods used.

D.2. Analysis of Benign Thyroid Nodule Risk

D.2.a. Primary Analysis

Two hundred forty-nine living evaluable participants had diagnoses of benign thyroid nodule(s) based on HTDS or prior histology or cytology. Fourteen of these cases were out-of-area participants, for whom the CIDER program could not calculate dose estimates. The number of cases and proportion with benign thyroid nodule(s) are shown by sex, dose category and basis for diagnosis in Table IX.D-10. The numbers and proportions of cases of additional disease outcomes related to benign thyroid nodule are shown in Table IX.D-11.

Table IX.D-10. Diagnoses of Benign Thyroid Nodule by Sex, Dose Category, and Basis for Diagnosis

A. Female

Thyroid Radiation Dose	Living Evaluable Female	Primary Definition: Cases Based on HTDS or Prior Histologic or Cytologic Diagnosis		1st Alternative Definition: Cases Based on HTDS or Prior Histology, Cytology, or Clinical Diagnosis		2nd Alternative Definition: Cases Based on Any Diagnosis or Participant or CATI Report	
(mGy)	No.	No.	%	No.	%	No.	%
Out of Area	125	10	8.0	12	9.6	13	10.4
< 10	182	20	11.0	24	13.2	25	13.7
10-49	320	31	9.7	34	10.6	34	10.6
50-99	313	27	8.6	31	9.9	31	9.9
100-149	220	19	8.6	21	9.5	23	10.5
150-199	126	17	13.5	18	14.3	19	15.1
200-299	139	15	10.8	17	12.2	19	13.7
300-399	144	12	8.3	16	11.1	16	11.1
400-999	171	19	11.1	20	11.7	20	11.7
1000+	7	0	--	0	--	0	--
Total	1747	170	9.7	193	11.0	200	11.4

B. Male

Thyroid Radiation Dose	Living Evaluable Male	Primary Definition: Cases Based on HTDS or Prior Histologic or Cytologic Diagnosis		1st Alternative Definition: Cases Based on HTDS or Prior Histology, Cytology, or Clinical Diagnosis		2nd Alternative Definition: Cases Based on Any Diagnosis or Participant or CATI Report	
(mGy)	No.	No.	%	No.	%	No.	%
Out of Area	124	4	3.2	4	3.2	4	3.2
< 10	186	7	3.8	8	4.3	8	4.3
10-49	314	19	6.1	19	6.1	19	6.1
50-99	310	14	4.5	23	7.4	24	7.7
100-149	171	7	4.1	9	5.3	9	5.3
150-199	109	6	5.5	6	5.5	6	5.5
200-299	148	13	8.8	14	9.5	14	9.5
300-399	160	5	3.1	6	3.8	6	3.8
400-999	154	3	1.9	4	2.6	6	3.9
1000+	17	1	5.9	1	5.9	1	5.9
Total	1693	79	4.7	94	5.6	97	5.7

Table IX.D-11. Additional Disease Outcomes Related to Benign Thyroid Nodule by Sex and Estimated Dose (cases based on primary definition of benign thyroid nodule, i.e., HTDS or prior histologic or cytologic diagnoses only)

A. Female

Thyroid Radiation Dose	Living Evaluable Female	Benign Thyroid Nodule or Nodule Suspicious for Follicular Neoplasm		Benign Thyroid Nodule Excluding Non-neoplastic Disease		Solitary Benign Thyroid Nodule Detected without Ultrasound		Benign Thyroid Nodule Excluding Colloid Only Nodules		Colloid Nodules	
(mGy)	No.	No.	%	No.	%	No.	%	No.	%	No.	%
OOA	125	11	8.8	8	6.4	5	4.0	9	7.2	7	5.6
< 10	182	23	12.6	14	7.7	5	2.7	18	9.9	14	7.7
10-49	320	32	10.0	24	7.5	15	4.7	27	8.4	25	7.8
50-99	313	30	9.6	15	4.8	13	4.2	26	8.3	20	6.4
100-149	220	21	9.5	15	6.8	3	1.4	18	8.2	17	7.7
150-199	126	18	14.3	12	9.5	7	5.6	17	13.5	13	10.3
200-299	139	15	10.8	11	7.9	4	2.9	14	10.1	10	7.2
300-399	144	13	9.0	6	4.2	5	3.5	12	8.3	12	8.3
400-999	171	21	12.3	12	7.0	7	4.1	17	9.9	18	10.5
1000+	7	0	--	0	--	0	--	0	--	0	--
Total	1747	184	10.5	117	6.7	64	3.7	158	9.0	136	7.8

B. Male

Thyroid Radiation Dose	Living Evaluable Female	Benign Thyroid Nodule or Nodule Suspicious for follicular neoplasm		Benign Thyroid Nodule Excluding Non-neoplastic Disease		Solitary Benign Nodule Detected without Ultrasound		Benign Thyroid Nodule Excluding Colloid Only Nodules		Colloid Nodules	
(mGy)	No.	No.	%	No.	%	No.	%	No.	%	No.	%
OOA	124	4	3.2	2	1.6	1	0.8	4	3.2	4	3.2
< 10	186	7	3.8	5	2.7	2	1.1	7	3.8	6	3.2
10-49	314	19	6.1	14	4.5	6	1.9	16	5.1	17	5.4
50-99	310	15	4.8	9	2.9	4	1.3	13	4.2	13	4.2
100-149	171	7	4.1	5	2.9	3	1.8	7	4.1	6	3.5
150-199	109	7	6.4	5	4.6	0	--	6	5.5	3	2.8
200-299	148	13	8.8	10	6.8	4	2.7	12	8.1	9	6.1
300-399	160	5	3.1	5	3.1	2	1.3	5	3.1	4	2.5
400-999	154	3	1.9	2	1.3	2	1.3	3	1.9	2	1.3
1000+	17	1	5.9	1	5.9	0	--	0	--	1	5.9
Total	1693	81	4.8	58	3.4	24	1.4	73	4.3	65	3.8

OOA = out of area participant

Parameter estimates for the linear dose-response model based on the 3191 in-area participants are shown in row 1 of Table IX.D-12 below. Based on maximum likelihood analysis of the sex-stratified linear probability model, and using the primary dose estimates, the estimated slope B was slightly less than zero (–0.008 per Gy) with Bonferroni-adjusted 95% CI ranging from less than –0.022 to 0.041 per Gy, providing no evidence that cumulative incidence increased with increasing dose (one-tailed p = 0.68). The corresponding estimated background rates for diagnosis of benign thyroid nodule were 0.100 with confidence interval (0.081, 0.119) for women and 0.049 with confidence interval (0.034, 0.064) for men. Very similar results were obtained when the model was fit by the method of least squares using ungrouped or grouped data (Table IX.D-12, rows 2 and 3).

Table IX.D-12. Summary of Dose-Response Results for Diagnoses of Benign Thyroid Nodule

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
1.	Primary definition (HTDS or prior, histologic or cytologic diagnosis)	Linear	Primary	None	MLE	.100 ± .008 (.081, .119)	.049 ± .006 (.034, .064)	–.008 ± .015 (< –.022, .041)	0.68
2.	Primary definition	Linear	Primary	None	LSU	.100 ± .007 (.082, .117)	.049 ± .008 (.031, .067)	–.006 ± .021 (–.055, .043)	0.61
3.	Primary definition	Linear	Primary	None	LSG	.101 ± .008 (.083, .119)	.050 ± .008 (.031, .069)	–.013 ± .024 (–.069, .044)	0.70
4.	Alternative def. #1 (HTDS or prior, histologic, cytologic, or clinical diagnosis)	Linear	Primary	None	MLE	.114 ± .008 (.094, .134)	.060 ± .007 (.044, .075)	–.013 ± .016 (< –.026, .037)	0.77
5.	Alternative def. #2 (Any diagnosis or participant/respondent report)	Linear	Primary	None	MLE	.117 ± .008 (.096, .137)	.061 ± .007 (.044, .077)	–.008 ± .018 (< –.027, .046)	0.67
6.	Benign thyroid nodule and nodules suspicious for follicular neoplasm	Linear	Primary	None	MLE	.108 ± .008 (.089, .128)	.050 ± .006 (.036, .065)	–.008 ± .015 (< –.022, .041)	0.69

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.D-12. Summary of Dose-Response Results for Diagnoses of Benign Thyroid Nodule (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
7.	Benign thyroid nodule excluding non-neoplastic disease	Linear	Primary	None	MLE	.068 ± .007 (.052, .084)	.036 ± .005 (.024, .049)	–.003 ± .013 (< –.016, .039)	0.60
8.	Solitary benign thyroid nodule detected without ultrasound	Linear	Primary	None	MLE	.037 ± .005 (.024, .050)	.015 ± .004 (.006, .025)	–.005 ± .014 (< –.006, .032)	0.63
9.	Benign thyroid nodule excluding colloid only nodules	Linear	Primary	None	MLE	.095 ± .009 (.075, .116)	.047 ± .006 (.031, .062)	–.019 ± .025 (NE, .026)	0.91
10.	Benign colloid nodules	Linear	Primary	None	MLE	.080 ± .007 (.062, .097)	.039 ± .005 (.026, .052)	–.002 ± .015 (< –.018, .044)	0.56
11.	Primary definition	LQ	Primary	None	LSU	.100 ± .008 (.080, .120)	.049 ± .008 (.029, .070)	Lin: –.009 ± .035 (–.096, .078) Quad: .003 ± .023 (–.055, .060)	Quad: 0.90

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.D-12. Summary of Dose-Response Results for Diagnoses of Benign Thyroid Nodule (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
12.	Primary definition	Logistic	Primary	None	MLE	.100 (.081, .123)	.048 (.036, .065)	–.092 ± .316 (–.849, .666)	0.62
13.	Primary definition	Linear	Primary	Exclude dose > 1000 mGy	MLE	.103 ± .009 (.082, .124)	.051 ± .007 (.035, .067)	–.021 ± .026 (< –.058, >.045)	0.79
14.	Primary definition	Linear	Primary	Exclude dose > 400 mGy	MLE	.098 ± .009 (.075, .120)	.051 ± .008 (.032, .070)	.001 ± .045 (–.102, .113)	0.49
15.	Primary definition	Linear	Primary	Exclude Ok and F/S geostrata	MLE	.098 ± .008 (.078, .118)	.047 ± .006 (.032, .062)	–.004 ± .017 (<–.021, .047)	0.60
16.	Primary definition	Linear	Alt. #1	None	MLE	.100 ± .008 (.081, .119)	.049 ± .006 (.034, .064)	–.007 ± .015 (< –.022, .039)	0.68
17.	Primary definition	Linear	Alt. #2	None	MLE	.101 ± .008 (.083, .119)	.050 ± .006 (.036, .065)	–.013 ± .010 (–.026, .023)	0.86
18.	Primary definition	Linear	Primary	Include OOA (scoping analysis #1)	MLE	.098 ± .008 (.080, .116)	.047 ± .006 (.034, .061)	–.004 ± .016 (< –.021, .045)	0.60
19.	Primary definition	Linear	Primary	Include OOA (scoping analysis #2)	MLE	.098 ± .008 (.080, .116)	.047 ± .006 (.034, .061)	–.005 ± .016 (< –.021, .044)	0.62

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

D.2.b. Alternative Definitions for Diagnosis of Benign Thyroid Nodule

Two alternative definitions for cases of benign thyroid nodule were considered. The first alternative added the 38 participants with HTDS or prior clinical diagnoses of benign thyroid nodule(s), for a total of 287 cases (271 in-area, and 16 out-of-area). The second added another 10 participants based solely on a report from the participant or his/her CATI respondent, for a total of 297 cases (280 in-area, and 17 out-of-area). As shown in rows 4 and 5 of Table IX.D-12 above, the parameter estimates for the linear dose-response model using either of these alternative definitions were essentially identical to those obtained in the primary analysis. In particular, the estimated slope of the linear dose-response model was less than zero for all three definitions of benign thyroid nodule, providing no evidence for any definition that the cumulative incidence of benign thyroid nodule increased with increasing dose ($p = 0.68$, 0.77, and 0.67 for the primary and first and second alternative definitions, respectively).

D.2.c. Additional Disease Outcomes Related to Benign Thyroid Nodule

D.2.c.1. Benign Thyroid Nodules and Nodules Suspicious for Follicular Neoplasm

Since most of the thyroid nodules classified as suspicious for follicular neoplasm were likely to be benign, the dose-response was also analyzed for the outcome of benign thyroid nodules and nodules suspicious for follicular neoplasm (Table IX.D-12, row 6). The estimated dose-response for this outcome was slightly negative (–0.008 per Gy with Bonferroni-adjusted 95% CI ranging from less than –0.022 to 0.041 per Gy), and consequently there was no evidence that the cumulative incidence of such nodules increased significantly with increasing dose ($p = 0.69$).

D.2.c.2 Benign Thyroid Nodule Excluding Non-neoplastic Disease

In order to investigate the possibility that a radiation-related increase in risk of benign thyroid nodules might be masked by the presence of nodules associated with other, nonradiogenic diseases, the dose-response was also analyzed for the outcome of benign thyroid nodule excluding non-neoplastic disease. The estimated dose-response for this outcome was also slightly negative (–0.003 per Gy with Bonferroni-adjusted 95% CI ranging from less than –0.016 to 0.039 per Gy, Table IX.D-12, row 7), and consequently there was no evidence that the cumulative incidence of such nodules increased significantly with increasing dose ($p = 0.60$).

D.2.c.3. Solitary Benign Thyroid Nodule Detected Without Ultrasound

As shown in row 8 of Table IX.D-12, the estimated slope of the dose-response for the outcome of solitary benign thyroid nodule detected without ultrasound was not significantly greater than zero (-0.005 per Gy, with Bonferroni-adjusted 95% confidence limits ranging from less than –0.006 to 0.032). Consequently there was no evidence that the cumulative incidence of such nodules increased significantly with increasing dose ($p = 0.63$).

D.2.c.4. Benign Thyroid Nodule Excluding Colloid-Only Nodules

The estimated slope of the dose-response for benign thyroid nodules excluding colloid-only nodules was slightly negative (–0.019 per Gy, Table IX.D-12, row 9). The Bonferroni-adjusted 95% lower confidence limit could not be estimated, and the upper confidence limit was 0.026 per Gy ($p = 0.91$).

D.2.c.5 Benign Colloid Nodules

The majority of participants with diagnoses of benign thyroid nodules had colloid nodules. Among the 3191 in-area participants, the cumulative incidence of colloid nodules did not increase significantly with increasing dose. As shown in row 10 of Table IX.D-12 above, the estimated slope was -0.002 per Gy, with Bonferroni-adjusted 95% confidence interval ranging from less than -0.018 to 0.044 per Gy (p=0.56).

D.2.d. Alternative Dose-Response Functions

As shown in row 11 of Table IX.D-12, the estimated regression coefficient for the dose-squared term in the linear-quadratic dose-response model [5] was 0.003 with Bonferroni-adjusted 95% confidence interval ranging from –0.055 to 0.060. Thus the addition of a quadratic term did not significantly improve the fit of the model (p = 0.90).

Parameter estimates for the sex-stratified logistic dose-response model [2] are shown in row 12 of Table IX.D-12. The estimated coefficient of radiation dose was less than zero (–0.092 per Gy, with Bonferroni-adjusted 95% confidence limits –0.849 and 0.666), providing no evidence that risk of benign thyroid nodule increased significantly with increasing dose (p = 0.62).

D.2.e. Effect of Excluding Participants in High Dose Categories

As rows 13 and 14 of Table IX.D-12 show, when participants in high dose categories were excluded, there was no evidence that the cumulative incidence of benign thyroid nodules increased with increasing dose.

D.2.f. Effect of Excluding Okanogan and Ferry/Stevens Geostrata

When participants in the Okanogan and Ferry/Stevens geostrata were excluded, the estimated slope B was not significantly greater than zero (-.004 per Gy, with Bonferroni-adjusted 95% CI ranging from less than -0.021 to 0.047 per Gy; Table IX.D-12, row 15), providing no evidence that the cumulative incidence of benign thyroid nodule increased with increasing dose (p=0.60).

D.2.g. Analysis of Benign Thyroid Nodules in Relation to Alternative Dose Estimates

As shown in rows 16 and 17 of Table IX.D-12, the cumulative incidence of benign thyroid nodule did not increase significantly in relation to either of the alternative dose estimates.

D.2.h. Scoping Analysis Regarding Out-of-Area Participants

See section VIII.C.1.a.3 for a description of the scoping analyses that were performed to assess the possible impact of including the 249 out-of-area participants. As summarized in rows 18 and 19 of Table IX.D-12, in both analyses the inclusion of the out-of-area participants had almost no effect on the estimated slope of the dose-response. In particular, the estimated slope of the dose-response was slightly negative in both scoping analyses, providing no evidence that cumulative incidence increased with increasing dose (p = 0.60 and 0.62 for the first and second scoping analysis, respectively).

D.2.i. Analysis of Benign Thyroid Nodule in Relation to Alternative Representations of Exposure

In the analyses by geostratum and by dichotomous exposure variable, the sex and age-adjusted comparisons of cumulative incidence were performed as described in section VIII.C.2.a.2.

D.2.i.1. Analysis by Geostratum

As shown in Table IX.D-13, among the entire 3440 living evaluable participants, the proportions with benign thyroid nodules ranged from 10/75 (13.3% in the Okanogan County geostratum) to 11/179 (6.1%, Richland) for women, and from 14/156 (9.0%, Adams County) to 2/76 (2.6%, Franklin County) for men ($p = 0.028$ for heterogeneity among the nine geostrata). In particular the percentages with benign thyroid nodules were somewhat higher in the Okanogan and Ferry/Stevens geostrata (11.9% for women, 6.0% for men) than in the remaining geostrata (9.5% and 4.6%, respectively; $p = 0.048$). Since it was likely that participants in the Okanogan and Ferry/Stevens geostrata tended to have lower thyroid doses from Hanford's ^{131}I than those in other geostrata, it does not appear that these differences can be attributed to an effect of Hanford's ^{131}I.

Table IX.D-13. Diagnoses of Benign Thyroid Nodule Based on HTDS or Prior Histologic or Cytologic Evidence, by Geostratum and Sex

	Female			Male			Total		
Geostratum	No.	Cases	%	No.	Cases	%	No.	Cases	%
Richland	179	11	6.1	173	7	4.0	352	18	5.1
Pasco/Kennewick	508	42	8.3	501	13	2.6	1009	55	5.5
Benton County	376	43	11.4	358	23	6.4	734	66	9.0
Franklin County	73	7	9.6	76	2	2.6	149	9	6.0
Adams County	165	18	10.9	156	14	9.0	321	32	10.0
Walla Walla (city)	133	13	9.8	131	5	3.8	264	18	6.8
Walla Walla County	170	19	11.2	164	7	4.3	334	26	7.8
Okanogan County	75	10	13.3	64	4	6.3	139	14	10.1
Ferry/Stevens Counties	68	7	10.3	70	4	5.7	138	11	8.0
Total	1747	170	9.7	1693	79	4.7	3440	249	7.2

D.2.i.2. Analysis by Dichotomous Exposure Variable

Of the 1257 participants included in these analyses, 102 (8.1%) had a diagnosis of benign thyroid nodule(s) based on an HTDS or prior histologic or cytologic examination (see Table IX.D-14). These included 53/580 (9.1%) in the high exposure group and 49/677 (7.2%) in the low exposure group. After adjusting for the effects of sex and age at HTDS clinic in the logistic regression analysis, there was no statistically significant evidence that the cumulative incidence of benign thyroid nodule was elevated in the high exposure group ($p = 0.20$).

Table IX.D-14. Diagnoses of Benign Thyroid Nodule based on HTDS or Prior Histologic or Cytologic Evidence, by Exposure Group and Sex

	Female			Male			Total		
Exposure Group	No.	Cases	%	No.	Cases	%	No.	Cases	%
Low	351	34	9.7	326	15	4.6	677	49	7.2
High	298	36	12.1	282	17	6.0	580	53	9.1
Total	649	70	10.8	608	32	5.3	1257	102	8.1

D.2.j. Confounding and Effect Modification

As described in section VIII above, additional sex-stratified logistic regression models were investigated to examine the possibility that confounding might influence the primary dose-response results, and to search for factors that might modify a radiation dose-response. These analyses were based on the primary definition of benign thyroid nodules, i.e., those with an HTDS or prior histologic or cytologic diagnosis, and on the primary dose estimates. Table IX.D-15 displays results for models including sex, age at first exposure to Hanford ^{131}I (prenatal, and < 180 days), age at HTDS examination, estimated dose from the NTS, history of any other cancer other than thyroid, and HTDS interview type (CATI versus expanded In-Person Interview).

Note that sex was not analyzed as a possible confounder since its effect was already adjusted for in the sex-stratified model. Adjusting for the possibility of confounding by any of the other covariates in Table IX.D-15 did not markedly change the estimated regression coefficient. For example, adjusting for a potential confounding effect of exposure to Hanford's ^{131}I in the HEDR domain before age 180 days changed the estimated coefficient from –0.092 to –0.121, a small change when considered in relation to the confidence intervals for these two estimates, (–0.849, 0.666) and (–0.966, 0.724), respectively. Moreover the adjusted estimate remained less than zero. Consequently, there was no evidence that a confounding effect of this age covariate obscured a positive dose-response for benign thyroid nodule. This pattern is evident for all of the covariates other than sex in Table IX.D-15.

The analyses of effect modification address the question of whether the dose-response might vary according to other characteristics of the study participants. This was tested by comparing the estimated regression coefficients for the groups defined by each covariate. As shown in Table IX.D-15, the regression coefficients did not differ significantly between the groups defined by any of the covariates, suggesting that none of them was a significant modifier of a radiation dose-response for benign thyroid nodule.

Table IX.D-15. Confounding and Effect Modification by Sex, Age at Exposure or HTDS Examination, Estimated Thyroid Dose from NTS, History of Any Cancer Other Than Thyroid and Interview Type: Benign Thyroid Nodule

		Estimated Dose-Response Coefficient (per Gy)				
Covariate (0=No, 1=Yes)	Yes / Total	Unadjusted	Adjusted for Confounding	Including Effect Modification Group 0	Including Effect Modification Group 1	P
Female?	1622 / 3191	–.092 ± .316 (–.849, .666)	Not Applicable	–.454 ± .601 (–1.96, 1.05)	.070 ± .368 (–.849, .990)	.45
Prenatal exposure?	1034 / 3191	–.092 ± .316 (–.849, .666)	–.165 ± .324 (–1.00, .670)	–.147 ± .367 (–1.11, .821)	–.230 ± .687 (–2.04, 1.58)	.91
1st exposure before age 180 days?	1478 / 3191	–.092 ± .316 (–.849, .666)	–.113 ± .327 (–.954, .728)	–.161 ± .522 (–1.54, 1.22)	–.082 ± .414 (–1.18, 1.01)	.91
Age at exam >50?	2001 / 3191	–.092 ± .316 (–.849, .666)	–.222 ± .333 (–1.08, .634)	–.516 ± .737 (–2.46, 1.43)	–.135 ± .374 (–1.12, .853)	.64
NTS ^{131}I dose > 5.3 mGy?	1566 / 3187	–.097 ± .318 (–.858, .665)	–.109 ± .326 (–.949, .731)	.153 ± .393 (–.884, 1.19)	–.588 ± .608 (–2.19, 1.02)	.29
History of any cancer other than thyroid?	248 / 3186	–.091 ± .316 (–.848, .666)	–.091 ± .317 (–.909, .726)	–.263 ± .365 (–1.23, .700)	–.483 ± .560 (–.994, 1.96)	.30
Expanded In-Person Interview?	1212 / 3191	–.092 ± .316 (–.849, .666)	–.007 ± .319 (–.828, .814)	–.135 ± .497 (–1.45, 1.18)	–.083 ± .403 (–.980, 1.15)	.73

Entries in the table are estimate ± standard error for the regression coefficient, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

Tables IX.D-16 and IX.D-17 display similar results from analyses including history of medical or dental x-ray exposure or occupational exposure as potential confounding or effect modifying factors. The estimates of the regression coefficient calculated with adjustment for confounding are all close to the unadjusted estimates. Moreover the adjusted estimates all remained less than zero. Thus there was no evidence that a confounding effect of any of these covariates obscured a positive dose-response for benign thyroid nodule.

There is no evidence of any statistically significant effect modification by any of the covariates in Tables IX.D-16 and IX.D-17, with two possible exceptions.

- The estimated dose-response coefficient was markedly negative (–2.44) for the 398 participants with histories of IVP, but not for the majority of participants without such histories (0.118 with confidence interval ranging from –0.704 to 0.941; p=0.036).

- The estimated dose-response coefficient was markedly negative (–3.13) for the 442 participants with histories of occupations that might have involved exposure to radioactive materials or x-rays, but not for the majority of participants without such histories (0.112 with confidence interval ranging from –0.738 to 0.963; p=0.023).

The statistical significance of these differences must be interpreted with caution due to the large number of such comparisons that were performed. Moreover, neither of these two covariates identified a group of participants with a significantly positive dose-response.

Table IX.D-16. Confounding and Effect Modification by History of Medical and Dental Radiation Exposures: Benign Thyroid Nodule

		Estimated Dose-Response Coefficient (per Gy)				
Ever Had: (0=No, 1=Yes)	Yes / Total	Unadjusted	Adjusted for Confounding	Including Effect Modification Group 0	Including Effect Modification Group 1	P
CAT scan of the upper body?	775 / 3149	−.120 ± .320 (−.885, .646)	−.119 ± .319 (−.940, .701)	.057 ± .325 (−.800, .913)	−1.17± .88 (−3.51, 1.16)	.18
Diagnostic x-rays of the head?	1191 / 3155	−.063 ± .315 (−.816, .691)	−.057 ± .314 (−.866, .751)	.188 ± .353 (−.743, 1.12)	−.650 ± .626 (−2.30, 1.00)	.23
Diagnostic x-rays of the neck?	966 / 3167	−.085 ± .316 (−.842, .672)	−.109 ± .316 (−.924, .706)	−.051 ± .428 (−1.18, 1.08)	−.178 ± .477 (−1.44, 1.08)	.84
Diagnostic x-rays of chest or upper body, including mammograms?	2821 / 3173	−.095 ± .317 (−.854, .664)	−.087 ± .317 (−.904, .730)	−.014 ± 1.23 (−3.27, 3.24)	−.092 ± .328 (−.959, .774)	.95
Diagnostic x-rays of the stomach or mid-back?	692 / 3120	−.118 ± .325 (−.896, .659)	−.123 ± .325 (−.960, .715)	−.260 ± .379 (−1.26, .739)	.348 ± .629 (−1.31, 2.01)	.43
Barium enema?	825 / 3159	−.098 ± .317 (−.856, .660)	−.097 ± .317 (−.912, .719)	−.196 ± .382 (−1.20, .812)	.149 ± .562 (−1.33, 1.63)	.62
Upper GI?	1146 / 3177	−.117 ± .320 (−.882, .648)	−.116 ± .320 (−.940, .708)	−.154 ± .364 (−.806, 1.11)	−.696 ± .607 (−2.30, .906)	.22
Intravenous pyelogram?	398 / 3157	−.095 ± .318 (−.856, .667)	−.084 ± .319 (−.904, .737)	.118 ± .312 (−.704, .941)	−2.44 ± 1.33 (−5.96, 1.07)	.036
Fluoroscopy of the upper body?	246 / 3161	−.071 ± .316 (−.828, .686)	−.074 ± .317 (−.890, .742)	.022 ± .318 (−.818, .862)	−1.48 ± 1.42 (−5.21, 2.26)	.26
Nuclear scan (excluding thyroid scan)?	217 / 3162	−.091 ± .317 (−.851, .668)	−.088 ± .317 (−.905, .729)	−.017 ± .319 (−.859, .824)	−1.46 ± 1.61 (−5.71, 2.79)	.34
Dental x-rays that did not usually include a lead shield over the neck area?	1648 / 3191	−.092 ± .316 (−.849, .666)	−.095 ± .317 (−.911, .720)	.143 ± .414 (−.949, 1.24)	−.380 ± .495 (−1.69, .927)	.41

Entries in the table are estimate ± standard error for the regression coefficient, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

Table IX.D-17. Confounding and Effect Modification by Occupational History: Benign Thyroid Nodule

Have You Ever Worked in Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted	Adjusted for Confounding	Including Effect Modification		P
				Group 0	Group 1	
Any metal industry?	238 / 3191	−.092 ± .316 (−.849, .666)	−.083 ± .316 (−.896, .730)	−.090 ± .321 (−.937, .757)	.144 ± 1.77 (−4.53, 4.82)	.90
Any nuclear facility?	371 / 3191	−.092 ± .316 (−.849, .666)	−.094 ± .320 (−.917, .729)	−.127 ± .351 (−1.05, .798)	.081 ± .783 (−1.98, 2.15)	.81
Any other industry or occupation where you may have been exposed to radioactive materials or x-rays?	442 / 3191	−.092 ± .316 (−.849, .666)	−.114 ± .319 (−.936, .708)	.112 ± .322 (−.738, .963)	−3.13 ± 1.69 (−7.59, 1.33)	.023
Any of the above industries or occupations?	892 / 3191	−.092 ± .316 (−.849, .666)	−.062 ± .316 (−.875, .751)	−.045 ± .366 (−1.01, .922)	−.108 ± .623 (−1.75, 1.54)	.93

Entries in the table are estimate ± standard error for the regression coefficient, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

Table IX.D-18 displays the results of analyses of possible confounding or effect modification by smoking variables. There was no evidence that the dose-response was significantly confounded by either smoking variable, or that there was a dose-response that differed significantly according to smoking history.

Table IX.D-18. Confounding and Effect Modification by Smoking: Benign Thyroid Nodule

Have You Ever Smoked Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted	Adjusted for Confounding	Including Effect Modification		P
				Group 0	Group 1	
Cigarettes (unfiltered or filtered)?	1854 / 3183	−.087 ± .316 (−.843, .668)	−.085 ± .316 (−.899, .729)	−.135 ± .533 (−1.54, 1.27)	−.057 ± .390 (−1.09, .972)	.91
Any of cigarettes, cigar or pipe?	1900 / 3183	−.087 ± .316 (−.843, .668)	−.085 ± .316 (−.898, .729)	−.034 ± .535 (−1.45, 1.38)	−.111 ± .394 (−1.15, .927)	.91

Entries in the table are estimate ± standard error for the regression coefficient, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

D.2.k. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for benign thyroid nodule are shown in Figure IX.D-1 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in that figure are calculated at the 98.33% confidence level, i.e., are adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates.

While the point estimate of the slope is greater than 0 for 30 of the 100 realizations, the confidence interval includes 0 for all 100 realizations. Also shown in Figure IX.D-1 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean and mean of each participant's 100 dose realizations. In summary, for none of the 100 realizations of the estimated doses was there a statistically significant dose-response, and for the majority of realizations the estimated slope was less than 0.

Figure IX.D-1. Plot of Estimated Slope and 95% CI by Dose Realization: Benign Thyroid Nodule

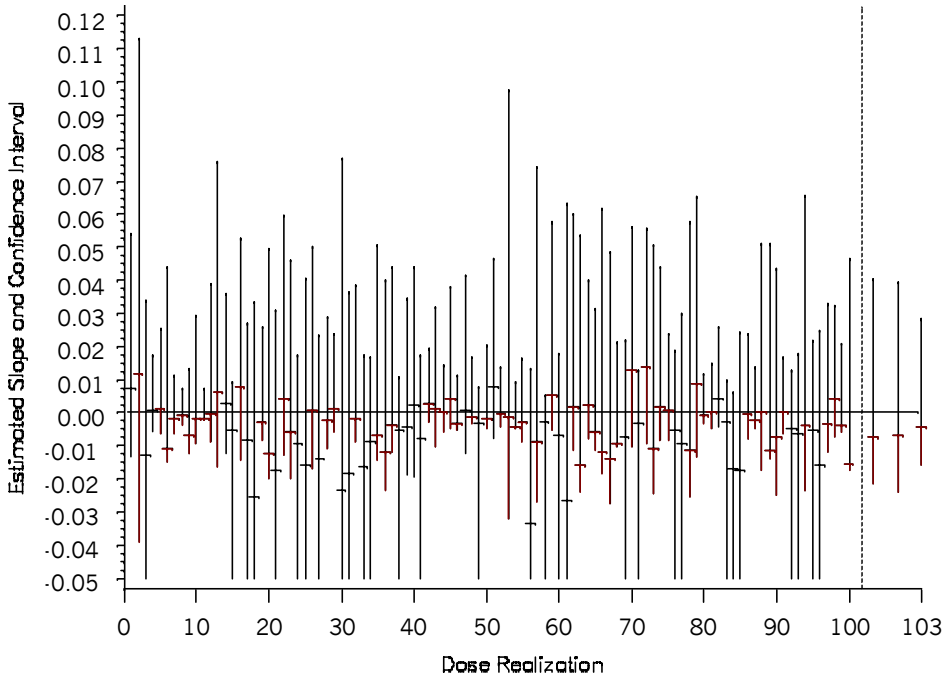
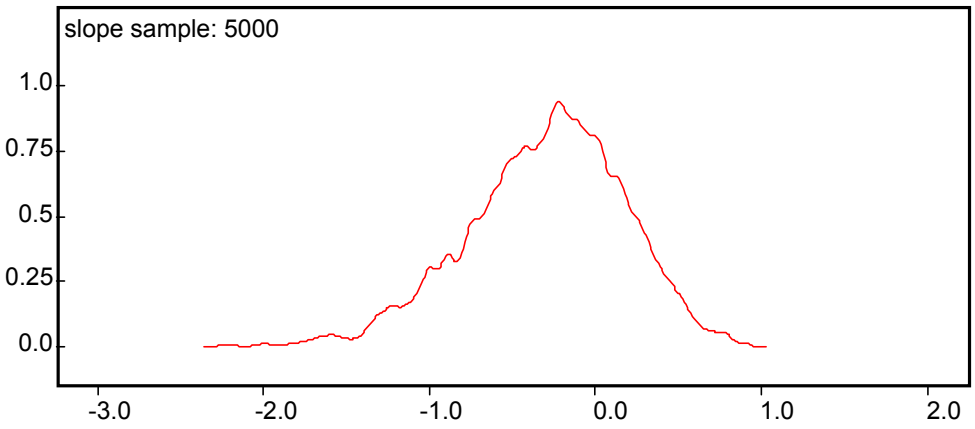


Figure IX.D-2 displays the distribution of the 5000 estimates of the logistic regression coefficient obtained by the simulation procedures described in section VIII.C.2.b.3 above. It is evident from the figure that most of the estimates were between about –2.0 and 1.0. The estimate was less than or equal to 0 for 3608 of the 5000 replications, implying an empirical one-tailed p-value of 0.72. The median estimate was –0.25, and the upper and lower percentiles corresponding to the Bonferroni-adjusted 95% confidence interval were –1.60 and 0.70. These may be compared to the estimate of –.092 with confidence interval (–.849, .666) obtained using the median dose estimates without adjustment for uncertainty. Thus, this method of adjusting the estimated logistic regression coefficient for the uncertainty in the dose estimates did not provide evidence that the cumulative incidence of benign thyroid nodule increased with increasing dose.

Figure IX.D-2. Distribution of Simulation Estimates of Logistic Regression Coefficient: Benign Thyroid Nodule



E. Total Thyroid Neoplasia

E.1. Occurrence of Total Thyroid Neoplasia

The outcome of total thyroid neoplasia was defined to include participants with thyroid carcinoma based on HTDS or prior histology or benign thyroid nodule with a histologic type of follicular adenoma, based on HTDS or prior histology.

Among the 3440 living evaluable participants 33 (1.0%) had a diagnosis of total thyroid neoplasia, with the percentage of cases slightly higher for women (1.1%) compared to men (0.8%) (Table IX.E-1).

Table IX.E-1. Total Thyroid Neoplasia, by Sex

Diagnosis of Thyroid Cancer or Follicular Adenoma	Female		Male		Total	
	No.	%	No.	%	No.	%
Yes	20	1.1	13	0.8	33	1.0
No	1727	98.9	1680	99.2	3407	99.0
Total	1747	100.0	1693	100.0	3440	100.0

E.2. Analysis of Total Thyroid Neoplasia Risk

E.2.a. Primary Analysis

The proportions of living evaluable participants with total thyroid neoplasia are shown by sex, in-area status, and dose group in Table IX.E-2.

Table IX.E-2. Diagnoses of Total Thyroid Neoplasia by Sex and Dose Category

Thyroid Radiation Dose (mGy)	Female			Male		
	Living Evaluable	Primary Definition: Cases Based on HTDS or Prior Histologic Diagnosis		Living Evaluable	Primary Definition: Cases Based on HTDS or Prior Histologic Diagnosis	
	No.	No.	%	No.	No.	%
Out of Area	125	2	1.6	124	3	2.4
< 10	182	2	1.1	186	2	1.1
10-49	320	4	1.3	314	3	1.0
50-99	313	4	1.3	310	0	--
100-149	220	2	0.9	171	0	--
150-199	126	2	1.6	109	2	1.8
200-299	139	2	1.4	148	1	0.7
300-399	144	1	0.7	160	1	0.6
400-999	171	1	0.6	154	0	--
1000+	7	0	--	17	1	5.9
Total	1747	20	1.1	1693	13	0.8

Parameter estimates for the linear dose-response model based on the 3191 in-area participants are shown in row 1 of Table IX.E-3 below. Based on maximum likelihood analysis of the sex-stratified linear probability model, the estimated slope B was 0.001 per Gy with Bonferroni-adjusted 95% CI ranging from less than –0.003 to 0.022 per Gy, providing no evidence that cumulative incidence increased with increasing dose (one-tailed p = 0.42). The corresponding estimated background rates for diagnosis of total thyroid neoplasia were 0.011 with confidence interval (0.004 , 0.018) for women and 0.006 with confidence interval (0.001, 0.012) for men. Very similar results were obtained when the model was fit by the method of least squares using ungrouped or grouped data (Table IX.E-3, rows 2 and 3).

Table IX.E-3 Dose-Response Results for Diagnoses of Total Thyroid Neoplasia

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
1.	Primary definition (HTDS or prior, histologic diagnosis)	Linear	Primary	None	MLE	.011 ± .003 (.004, .018)	.006 ± .002 (.001, .012)	.001 ± .006 (< –.003, .022)	0.42
2.	Primary definition	Linear	Primary	None	LSU	.011 ± .003 (.005, .017)	006 ± .003 (0*, .013)	.000 ± .007 (–.017, .018)	0.48
3.	Primary definition	Linear	Primary	None	LSG	.012 ± .003 (.006, .019)	.007 ± .003 (.001, .014)	–.006 ± .009 (–.026, .015)	0.75
4.	Primary definition	LQ	Primary	None	LSU	.011 ± .003 (.004, .019)	.007 ± .003 (0*, .014)	Lin: –.003 ± .013 (–.034, .028) Quad: .003 ± .008 (–.018, .023)	Quad: 0.74

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model. "0*" indicates that the lower confidence limit for a background rate was less than 0.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.E-3 Dose-Response Results for Diagnoses of Total Thyroid Neoplasia (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
5.	Primary definition	Logistic	Primary	None	MLE	.011 (.006, .021)	.006 (.003, .014)	.050 ± .833 (–1.94, 2.04)	0.48
6.	Primary definition	Linear	Primary	Exclude dose > 1000 mGy	MLE	.012 ± .003 (.005, .020)	.007 ± .002 (.001, .012)	-.006 ± .007 (< –.007, >.017)	0.77
7.	Primary definition	Linear	Primary	Exclude dose > 400 mGy	MLE	.012 ± .003 (.004, .020)	.007 ± .003 (0*, .013)	-.001 ± .015 (–.034, .040)	0.53
8.	Primary definition	Linear	Primary	Exclude OK and F/S geostrata	MLE	.010 ± .003 (.003, .016)	.006 ± .002 (.000, .012)	.002 ± .007 (<–.003, .023)	0.37

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model. "0*" indicates that the lower confidence limit for a background rate was less than 0.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.E-3 Dose-Response Results for Diagnoses of Total Thyroid Neoplasia (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
9.	Primary definition	Linear	Alt. #1	None	MLE	.012 ± .003 (.004, .019)	007 ± .003 (.000, .013)	-.003 ± .009 (NE, .015)	0.77
10.	Primary definition	Linear	Alt. #2	None	MLE	.011 ± .003 (.004, .018)	007 ± .003 (.001, .013)	-.003 ± .010 (NE, .010)	0.85
11.	Primary definition	Linear	Primary	Include OOA (scoping analysis #1)	MLE	.012 ± .003 (.005, .018)	.008 ± .002 (.002, .014)	-.001 ± .006 (<-.003, .019)	0.55
12.	Primary definition	Linear	Primary	Include OOA (scoping analysis #2)	MLE	.012 ± .003 (.005, .018	.008 ± .002 (.002, .014)	-.001 ± .006 (<-.003, >.018)	0.58

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model. "0*" indicates that the lower confidence limit for a background rate was less than 0.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, OK = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

E.2.b. Alternative Dose-Response Functions

As shown in row 4 of Table IX.E-3, the estimated regression coefficient for the dose-squared term in the linear-quadratic dose-response model [5] was 0.003 with Bonferroni-adjusted 95% confidence interval ranging from –0.018 to 0.023. Thus the addition of a quadratic term did not significantly improve the fit of the model ($p = 0.74$).

The regression parameter for the effect of dose in the sex-stratified logistic regression model [2] was estimated as 0.050 with Bonferroni-adjusted 95% CI ranging from –1.94 to 2.04 (Table IX.E-3, row 5). Thus there was no evidence from the logistic regression model that cumulative incidence of total thyroid neoplasia increased significantly with increasing dose ($p = 0.48$).

E.2.c. Effect of Excluding Participants in High Dose Categories

As shown in rows 6 and 7 of Table IX.E-3, excluding participants with doses above 1000 mGy or above 400 mGy resulted in slightly negative estimates for the slope of the dose-response, thus providing no evidence that risk increased with increasing dose ($p = 0.77$ and 0.53 based on participants with doses $\leq$ 1000 mGy and $\leq$ 400 mGy, respectively).

E.2.d. Effect of Excluding Okanogan and Ferry/Stevens Geostrata

As shown in row 8 of Table IX.E-3, if participants from the Okanogan and Ferry/Stevens geostrata are excluded, the estimated slope of the dose-response is 0.002 per Gy, with Bonferroni-adjusted 95% confidence interval ranging from less than - 0.003 to 0.023 per Gy; $p = 0.37$).

E.2.e. Analysis of Total Thyroid Neoplasia in Relation to Alternative Dose Estimates

As shown in rows 9 and 10 of Table IX.E-3, the cumulative incidence of total thyroid neoplasia did not increase significantly in relation to either of the alternative dose estimates.

E.2.f. Scoping Analysis Regarding Out-of-Area Participants

See section VIII.C.1.a.3 for a description of the scoping analyses that were performed to assess the possible impact of including the 249 out-of-area participants. As summarized in rows 11 and 12 of Table IX.E-3, in both analyses the inclusion of the out-of-area participants slightly decreased the estimated slope of the dose-response, but did not materially change the dose-response results.

E.2.g. Analysis of Total Thyroid Neoplasia in Relation to Alternative Representations of Exposure

In the analyses by geostratum and by dichotomous exposure variable, the sex and age-adjusted comparisons of cumulative incidence were performed as described in section VIII.C.2.a.2.

E.2.g.1. Analysis by Geostratum

As shown in Table IX.E-4, among the entire 3440 living evaluable participants, the proportions with thyroid neoplasia ranged from 0/73 (0% in the Franklin County geostratum) to 2/68 (2.9%, Ferry/Stevens Counties) for women, and from 0/76 (0%, Franklin County) to 1/64 (1.6%, Okanogan County) for men ($p = 0.41$ for heterogeneity among the nine geostrata). In particular the percentages with thyroid neoplasia were somewhat higher in the Okanogan and Ferry/Stevens geostrata (2.8% for women, 1.5% for men) than in the remaining geostrata (1.0% and 0.7%, respectively; $p = 0.037$). Since it was likely that participants in the Okanogan and Ferry/Stevens geostrata tended to have lower thyroid doses from Hanford's ^{131}I than those in other geostrata, it does not appear that these differences can be attributed to an effect of Hanford's ^{131}I.

Table IX.E-4. Diagnoses of Total Thyroid Neoplasia Based On Histologic or Cytologic Evidence from or Prior to the HTDS, by Geostratum and Sex

	Female			Male			Total		
Geostratum	No.	Cases	%	No.	Cases	%	No.	Cases	%
Richland	179	2	1.1	173	1	0.6	352	3	0.9
Pasco/Kennewick	508	4	0.8	501	3	0.6	1009	7	0.7
Benton County	376	5	1.3	358	4	1.1	734	9	1.2
Franklin County	73	0	--	76	0	--	149	0	--
Adams County	165	2	1.2	156	1	0.6	321	3	0.9
Walla Walla (city)	133	1	0.8	131	1	0.8	264	2	0.8
Walla Walla County	170	2	1.2	164	1	0.6	334	3	0.9
Okanogan County	75	2	2.7	64	1	1.6	139	3	2.2
Ferry/Stevens Counties	68	2	2.9	70	1	1.4	138	3	2.2
Total	1747	20	1.1	1693	13	0.8	3440	33	1.0

E.2.g.2. Analysis by Dichotomous Exposure Variable

Of the 1257 participants included in these analyses, 16 (1.3%) had a diagnosis of total thyroid neoplasia based on an HTDS or prior histologic or cytologic examination (see Table IX.E-5). These included 5/580 (0.9%) in the high exposure group and 11/677 (1.6%) in the low exposure group. After adjusting for the effects of sex and age at HTDS clinic in the logistic regression analysis, there was no statistically significant evidence that the cumulative incidence of total thyroid neoplasia was elevated in the high exposure group ($p = 0.73$).

Table IX.E-5. Diagnoses of Total Thyroid Neoplasia Based on HTDS or Prior Histologic or Cytologic Evidence, by Exposure Group and Sex

	Female			Male			Total		
Exposure Group	No.	Cases	%	No.	Cases	%	No.	Cases	%
Low	351	6	1.7	326	5	1.5	677	11	1.6
High	298	3	1.0	282	2	0.7	580	5	0.9
Total	649	9	1.4	608	7	1.2	1257	16	1.3

E.2.h. Confounding and Effect Modification

There were too few participants with diagnoses in the category of total thyroid neoplasia to warrant any analysis of confounding or effect modification.

E.2.i. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for total thyroid neoplasia are shown in Figure IX.E-1 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in that figure are calculated at the 98.33% confidence level, i.e., are adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates. While the point estimate of the slope is greater than 0 for 47 of the 100 realizations, the confidence interval includes 0 for all 100 realizations. Also shown in Figure IX.E-1 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean and mean of each participant's 100 dose realizations dose. In summary, for none of the 100 realizations of the estimated doses was there a statistically significant dose-response, and for about half of the realizations the estimated slope was less than 0.

Figure IX.E.1 Plot of Estimated Slope and 95% CI by Dose Realization: Total Thyroid Neoplasia

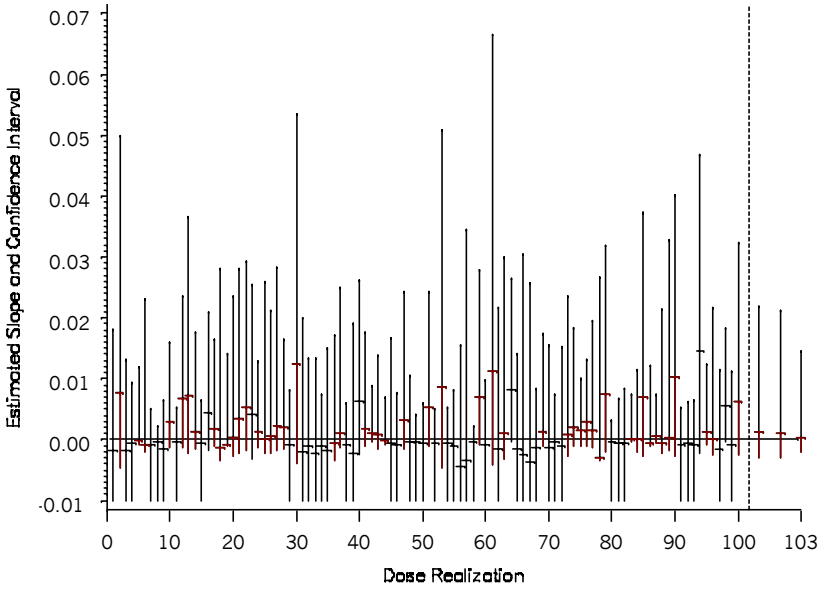
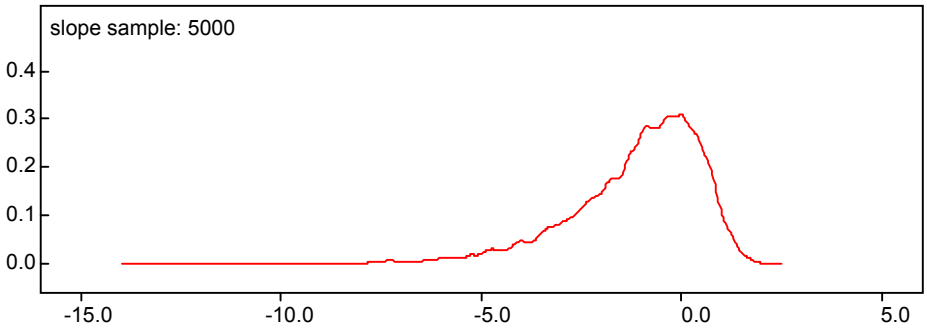


Figure IX.E-2 displays the distribution of the 5000 estimates of the logistic regression coefficient obtained by the simulation procedures described in section VIII.C.2.b.3 above. It is evident from the figure that most of the estimates were between about –6.0 and 1.5. The estimate was less than or equal to 0 for 3640 of the 5000 replications, implying an empirical one-tailed p-value of 0.73. The median estimate was –0.73, and the upper and lower percentiles corresponding to the Bonferroni-adjusted 95% confidence interval were –6.78 and 1.42. These may be compared to the estimate of 0.050 with confidence interval (–1.94, 2.04) obtained using the median dose estimates without adjustment for uncertainty. Thus this method of adjusting the estimated logistic regression coefficient for the uncertainty in the dose estimates did not provide evidence that the cumulative incidence of thyroid neoplasia increased with increasing dose.

Figure IX.E-2. Distribution of Simulation Estimates of Logistic Regression Coefficient: Total Thyroid Neoplasia



F. Any Thyroid Nodule

F.1. Occurrence of Any Thyroid Nodule

The primary and alternative definitions for the outcome of any thyroid nodule were as follows:

- Primary definition: HTDS or prior, histologic or cytologic diagnosis (281 cases)
- Alternative definition #1: HTDS or prior, histologic, cytologic or clinical diagnosis (320 cases)
- Alternative definition #2: Any diagnosis or participant/respondent report (330 cases).

The outcome of any thyroid nodule was defined by the presence of a diagnosis of one or more of benign thyroid nodule, thyroid carcinoma, or nodule suspicious for follicular neoplasm. Table IX.F-1 shows that 281 (8.2%) living evaluable participants had a diagnosis of any thyroid nodule based on histologic or cytologic evidence from the HTDS or prior, with the percentage about twice as high for women (11.0%) as for men (5.2%). Another 1.1% was based on clinical diagnoses by the HTDS or prior. There were 10 living evaluable participants with a diagnosis of any thyroid nodule based on reports from the participant or his/her CATI respondent.

Table IX.F-1. Basis for Diagnosis of Any Thyroid Nodule Disease, by Sex

	Female		Male		Total	
Diagnosis of Any Thyroid Nodule	No.	%	No.	%	No.	%
Yes	224	12.8	106	6.3	330	9.6
• Histologic diagnosis: HTDS	11	0.6	12	0.7	23	0.7
• Cytologic diagnosis: HTDS	156	8.9	67	4.0	223	6.5
• Prior histologic diagnosis	24	1.4	8	0.5	32	0.9
• Prior cytologic diagnosis	2	0.1	1	0.1	3	0.1
• Clinical diagnosis: HTDS	16	0.9	13	0.8	29	0.8
• Clinical diagnosis: prior	8	0.5	2	0.1	10	0.3
• Participant/respondent report	7	0.4	3	0.2	10	0.3
No	1521	87.1	1586	93.7	3107	90.3
Unknown	2	0.1	1	0.1	3	0.1
Total	1747	100.0	1693	100.0	3440	100.0

F.1.a. Additional Disease Outcomes Related to Any Thyroid Nodule

F.1.a.1. Any Solitary Thyroid Nodule Detected without Ultrasound

The outcome of any palpable solitary thyroid nodule detected without ultrasound was defined in order to simulate the effect of screening for thyroid disease by palpation only, i.e., without ultrasound examination. This analysis allows us to compare the prevalence of thyroid nodularity with older studies (e.g. The Framingham Study) that used only palpation to determine the prevalence of nodular thyroid disease. In HTDS a total of 117 living evaluable participants (83 women, 34 men) had diagnoses of such nodules (Table IX.F-2).

Table IX.F-2. Any Solitary Thyroid Nodule Detected without Ultrasound, by Sex

Any Solitary Thyroid Nodule	Female		Male		Total	
Detected without Ultrasound	No.	%	No.	%	No.	%
Yes	83	4.8	34	2.0	117	3.4
No	1664	95.2	1659	98.0	3323	96.6
Total	1747	100.0	1693	100.0	3440	100.0

For the majority of the 117 living evaluable participants with solitary thyroid nodules that were detected without ultrasound, i.e., by palpation, those nodules were also observed on the ultrasound examination. However for 21 (18%) of the 117, those nodules were not detected by ultrasound. Twelve (57%) of these 21 participants each had 1-6 discrete focal ultrasound abnormalities in addition to the palpable nodule which was not detected on ultrasound. In addition, 15 of 21 (71%) had documented Hashimoto's thyroiditis. Only 4 participants (0.1% of the 3429 living evaluable participants whose thyroid glands were visible in their ultrasound examinations) had a palpable nodule with a completely normal ultrasound scan. These results suggest that the reason for the discordance between palpation and ultrasound in this small group was the abnormal thyroid tissue that is present throughout the gland in individuals with Hashimoto's thyroiditis, a fact well known in clinical practice. Since only 4 participants had true palpable nodules that were not detected by ultrasound, a dose-response analysis of this specific outcome was not feasible.

F.2. Analysis of Any Thyroid Nodule Risk

F.2.a. Primary Analysis

The proportions with any thyroid nodule are shown by sex, in-area status, and dose group in Table IX.F-3 below. The numbers and proportions with diagnoses of any solitary thyroid nodule detected without ultrasound are also shown.

Table IX.F-3. Diagnoses of Any Thyroid Nodule by Sex, Estimated Dose, and Basis for Diagnosis

A. Female

Thyroid Radiation Dose (mGy)	Living Evaluable Female	Primary Definition: Cases Based on HTDS or Prior Histologic or Cytologic Diagnosis		1st Alternative Definition:Cases Based on HTDS or Prior Histology, Cytology, or Clinical Diagnosis		2nd Alternative Definition: Cases Based on Any Diagnosis or Participant or CATI Report		Any Solitary Thyroid Nodule Detected without Ultrasound	
	No.	No.	%	No.	No.	%	%	No.	%
OOA	125	13	10.4	16	12.8	17	13.6	7	5.6
< 10	182	24	13.2	28	15.4	29	15.9	8	4.4
10-49	320	34	10.6	37	11.6	37	11.6	18	5.6
50-99	313	31	9.9	35	11.2	35	11.2	16	5.1
100-149	220	22	10.0	24	10.9	26	11.8	5	2.3
150-199	126	19	15.1	20	15.9	21	16.7	8	6.3
200-299	139	15	10.8	17	12.2	19	13.7	4	2.9
300-399	144	14	9.7	18	12.5	18	12.5	7	4.9
400-999	171	21	12.3	22	12.9	22	12.9	10	5.8
1000+	7	0	--	0	--	0	--	0	--
Total	1747	193	11.0	217	12.4	224	12.8	83	4.8

B. Male

Thyroid Radiation Dose (mGy)	Living Evaluable Male	Primary Definition: Cases Based on HTDS or Prior Histologic or Cytologic Diagnosis		1st Alternative Definition:Cases Based on HTDS or Prior Histology, Cytology, or Clinical Diagnosis		2nd Alternative Definition: Cases Based on Any Diagnosis or Participant or CATI Report		Any Solitary Thyroid Nodule Detected without Ultrasound	
	No.	No.	%	No.	No.	%	%	No.	%
OOA	124	7	5.6	7	5.6	7	5.6	3	2.4
< 10	186	8	4.3	9	4.8	9	4.8	3	1.6
10-49	314	21	6.7	21	6.7	21	6.7	7	2.2
50-99	310	15	4.8	24	7.7	25	8.1	6	1.9
100-149	171	7	4.1	9	5.3	9	5.3	3	1.8
150-199	109	7	6.4	7	6.4	7	6.4	1	0.9
200-299	148	13	8.8	14	9.5	14	9.5	6	4.1
300-399	160	5	3.1	6	3.8	6	3.8	2	1.3
400-999	154	3	1.9	4	2.6	6	3.9	3	1.9
1000+	17	2	11.8	2	11.8	2	11.8	0	--
Total	1693	88	5.2	103	6.1	106	6.3	34	2.0

OOA = out of area

Parameter estimates for the linear dose-response model based on the 3191 in-area participants are shown in row 1 of Table IX.F-4 below. Based on maximum likelihood analysis of the sex-stratified linear probability model, the estimated slope B was slightly less than zero (–0.007 per Gy) with Bonferroni-adjusted 95% CI ranging from less than –0.023 to 0.043 per Gy, providing no evidence that cumulative incidence increased with increasing dose (one-tailed p = 0.65). The corresponding estimated background rates for outcome of any thyroid nodule were 0.112 with confidence interval (0.092, 0.132) for women and 0.053 with confidence interval (0.038 to 0.068) for men. Very similar results were obtained when the model was fit by least squares using ungrouped or grouped data (Table IX.F-4, rows 2 and 3).

Table IX.F-4. Dose-Response Results for Diagnoses of Any Thyroid Nodule

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
1.	Primary definition (HTDS or prior, histologic or cytologic diagnosis)	Linear	Primary	None	MLE	.112 ± .008 (.092, .132)	.053 ± .006 (.038, .068)	-.007 ± .016 (<-.023, .043)	0.65
2.	Primary definition	Linear	Primary	None	LSU	.112 ± .008 (.094, .131)	.053 ± .008 (.034, .072)	–.006 ± .022 (–.058, .045)	0.61
3.	Primary definition	Linear	Primary	None	LSG	.114 ± .008 (.095, .133)	.055 ± .008 (.035, .074)	–.017 ± .025 (–.078, .043)	0.75
4.	Alternative def. #1 (HTDS or prior, histologic, cytologic, or clinical diagnosis)	Linear	Primary	None	MLE	.126 ± .009 (.105, .147)	.063 ± .007 (.047, .080)	-.012 ± .017 (<-.028, .039)	0.75
5.	Alternative def. #2 (Any diagnosis or participant/respondent report)	Linear	Primary	None	MLE	.129 ± .009 (.108, .150)	.064 ± .007 (.048, .081)	-.007 ± .019 (<-.029, .047)	0.65

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value,"NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.F-4. Dose-Response Results for Diagnoses of Any Thyroid Nodule (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
6.	Solitary thyroid nodule detected without ultrasound	Linear	Primary	None	MLE	.047 ± .006 (.032, .061)	.020 ± .004 (.009, .030)	.001 ± .015 (<-.009, .042)	0.46
7.	Primary definition	LQ	Primary	None	LSU	.113 ± .008 (.092, .134)	.054 ± .009 (.032, .075)	Lin: –.014 ± .037 (–.105, .078) Quad: .006 ± .024 (–.054, .066)	Quad: 0.80
8.	Primary definition	Logistic	Primary	None	MLE	.112 (.092, .137)	.052 (.039, .070)	-.09 ± .30 (-.81, .63)	0.62
9.	Primary definition	Linear	Primary	Exclude dose > 1000 mGy	MLE	.117 ± .009 (.095, .140)	.055 ± .007 (.039, .072)	-.032 ± .026 (<-.062, >.035	0.88
10.	Primary definition	Linear	Primary	Exclude dose > 400 mGy	MLE	.112 ± .010 (.089, .135)	.056 ± .008 (.037, .076)	-.017 ± .046 (-.122, .099)	0.64
11.	Primary definition	Linear	Primary	Exclude Ok and F/S geostrata	MLE	.108 ± .009 (.087, .129)	.051 ± .007 (.035, .067)	-.003 ± .018 (<-.023, .050)	0.56
12.	Primary definition	Linear	Alt. #1	None	MLE	.113 ± .008 (.093, .133)	.054 ± .006 (.038, .069)	-.010 ± .015 (<-.024, .036)	0.74

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value,"NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.F-4. Dose-Response Results for Diagnoses of Any Thyroid Nodule (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
13.	Primary definition	Linear	Alt. #2	None	MLE	.114 ± .008 (.094, .133)	.055 ± .006 (.040, .069)	-.015 ± .010 (-.028, .021)	0.88
14.	Primary definition	Linear	Primary	Include OOA (scoping analysis #1)	MLE	.111 ± .008 (.092, .131)	.053 ± .006 (.038, .067)	-.006 ± .016 (<-.023, .043)	0.64
15.	Primary definition	Linear	Primary	Include OOA (scoping analysis #2)	MLE	.112 ± .008 (.093, .131)	.053 ± .006 (.039, .068)	-.007 ± .016 (<-.023, >.041)	0.66

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value,"NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

F.2.b. Alternative Definitions for Diagnosis of Any Thyroid Nodule

Two alternative definitions for cases of any thyroid nodule were considered. The first alternative added the 39 participants with HTDS or prior clinical diagnoses of any thyroid nodule, for a total of 320 (297 in-area, 23 out-of-area) cases. The second alternative criterion for defining cases of nodular disease added another 10 participants based solely on a report from the participant or his/her CATI respondent, for a total of 330 (306 in-area, 24 out-of-area) cases. As shown in rows 4 and 5 of Table IX.F-4 above, there was no evidence that the cumulative incidence of any thyroid nodule increased with increasing dose for either of these alternative definitions.

F.2.c. Additional Disease Outcomes Related to Any Thyroid Nodule

F.2.c.1. Any Solitary Thyroid Nodule Detected Without Ultrasound

As shown in row 6 of Table IX.F-4, the estimated slope of the dose-response for the outcome of any solitary thyroid nodule detected without ultrasound was not significantly greater than zero (0.001 per Gy with Bonferroni-adjusted 95% CI ranging from less than –0.009 to 0.042 per Gy). Consequently there was no evidence that the cumulative incidence of such nodules increased significantly with increasing dose ($p = 0.46$).

F.2.d. Alternative Dose-Response Functions

As shown in row 7 of Table IX.F-4, the estimated regression coefficient for the dose-squared term in the linear-quadratic dose-response model [5] was 0.006 with Bonferroni-adjusted 95% confidence interval ranging from –0.054 to 0.066. Thus the addition of a quadratic term did not significantly improve the fit of the model ($p = 0.80$).

In the analysis of any thyroid nodule based on HTDS or prior histologic or cytologic evidence, i.e., the primary criterion for defining cases with any thyroid nodule, the regression parameter for the effect of dose in the sex-stratified logistic regression model [2] was estimated as –0.09 with Bonferroni-adjusted 95% confidence interval ranging from –0.81 to 0.63 (Table IX.F-4, row 8). Thus there was no evidence from the logistic regression model that cumulative incidence of any thyroid nodule increased with increasing dose ($p = 0.62$).

F.2.e. Effect of Excluding Participants in High Dose Categories

As shown in rows 9 and 10 of Table IX.F-4, when participants in high dose categories were excluded, there was no evidence that the cumulative incidence of any thyroid nodule increased with increasing dose.

F.2.f. Effect of Excluding Okanogan and Ferry/Stevens Geostrata

Excluding Okanogan and Ferry/Stevens geostrata had no effect on the dose-response results, namely, there was no evidence that cumulative incidence of any thyroid nodule increased with increasing dose ($p = 0.56$; Table IX.F-4, row 11).

F.2.g. Analysis of Any Thyroid Nodule in Relation to Alternative Dose Estimates

When the first set of alternative dose estimates were used, the estimated slope B was -0.010 per Gy with Bonferroni-adjusted 95% confidence interval ranging from less than –0.024 to 0.036 (Table IX.F-4, row 12). For the second set of alternative dose estimates the estimated slope B was –0.015 per Gy with Bonferroni adjusted 95% confidence interval ranging from less than –0.028 to 0.021 (Table IX.F-4, row 13). Thus, for neither set of alternative dose estimates was there any evidence that the cumulative incidence of any thyroid nodule increased with increasing dose ($p = 0.74$ and 0.88 for the first and second set of alternative dose estimates, respectively).

F.2.h. Scoping Analysis Regarding Out-of-Area Participants

See section VIII.C.1.a.3 for a description of the scoping analyses that were performed to assess the possible impact of including the 249 out-of-area participants. As shown in rows 14 and 15 of Table IX.F-4, for neither scoping analysis was there any evidence that the cumulative incidence of any thyroid nodule increased with increasing dose ($p = 0.64$ and 0.66 for the first and second scoping analyses, respectively).

F.2.i. Analysis of Any Thyroid Nodule in Relation to Alternative Representations of Exposure

In both the analyses by geostratum and by dichotomous exposure variable, the sex and age-adjusted comparisons of cumulative incidence were performed as described in section VIII.C.2.a.2.

F.2.i.1. Analysis by Geostratum

As shown in Table IX.F-5, among the entire 3440 living evaluable participants, the proportions with any thyroid nodule (based on an HTDS or prior histologic or cytologic diagnosis) ranged from 12/75 (16.0% in the Okanogan County geostratum) to 14/179 (7.8%, Richland) for women, and from 14/156 (9.0%, Adams County) to 2/76 (2.6%, Franklin County) for men ($p = 0.032$ for heterogeneity among the nine geostrata). In particular the percentages with any thyroid nodule were somewhat higher in the Okanogan and Ferry/Stevens geostrata (15.4% for women, 6.7% for men) than in the remaining geostrata (10.7% and 5.1%, respectively; $p = 0.010$). Since it was likely that participants in the Okanogan and Ferry/Stevens geostrata tended to have lower thyroid doses from Hanford's ^{131}I than those in other geostrata, it does not appear that these differences can be attributed to an effect of Hanford's ^{131}I.

Table IX.F-5. Diagnoses of Any Thyroid Nodule with at Least One Outcome Based On Histologic or Cytologic Evidence from or Prior to the HTDS

	Female			Male			Total		
Geostratum	No.	Cases	%	No.	Cases	%	No.	Cases	%
Richland	179	14	7.8	173	8	4.6	352	22	6.3
Pasco/Kennewick	508	49	9.6	501	16	3.2	1009	65	6.4
Benton County	376	45	12.0	358	25	7.0	734	70	9.5
Franklin County	73	7	9.6	76	2	2.6	149	9	6.0
Adams County	165	20	12.1	156	14	9.0	321	34	10.6
Walla Walla (city)	133	15	11.3	131	6	4.6	264	21	8.0
Walla Walla County	170	21	12.4	164	8	4.9	334	29	8.7
Okanogan County	75	12	16.0	64	5	7.8	139	17	12.2
Ferry/Stevens Counties	68	10	14.7	70	4	5.7	138	14	10.1
Total	1747	193	11.0	1693	88	5.2	3440	281	8.2

F.2.i.2. Analysis by Dichotomous Exposure Variable

A total of 118 (9.4%) of the 1257 participants included in these analyses had a diagnosis of any thyroid nodule based on an HTDS or prior histologic or cytologic examination (see Table IX.F-6). These included 57/580 (9.8%) in the high exposure group and 61/677 (9.0%) in the low exposure group. After adjusting for the effects of sex and age at HTDS clinic in the logistic regression analysis, there was no evidence of greater cumulative incidence of any thyroid nodule in the high exposure group ($p = 0.38$).

Table IX.F-6. Diagnoses of Any Thyroid Nodule Based on HTDS or Prior Histologic or Cytologic Evidence, by Exposure Group and Sex

	Female			Male			Total		
Exposure Group	No.	Cases	%	No.	Cases	%	No.	Cases	%
Low	351	42	12.0	326	19	5.8	677	61	9.0
High	298	39	13.1	282	18	6.4	580	57	9.8
Total	649	81	12.5	608	37	6.1	1257	118	9.4

F.2.j. Confounding and Effect Modification

As described in section VIII above, additional sex-stratified logistic regression models were investigated to examine the possibility that the primary dose-response results might be influenced by confounding, and to search for factors that might modify a radiation dose-response. These analyses were based on the primary definition of any thyroid nodule, i.e., those with an HTDS or prior histologic or cytologic diagnosis, and on the primary dose estimates. Table IX.F-7 displays results for models including sex, age at first exposure to Hanford ^{131}I (prenatal, or < 180 days), age at HTDS examination, estimated dose from the NTS, history of any cancer other than thyroid, and HTDS interview type.

Note that sex was not analyzed as a possible confounder since its effect was already adjusted for in the sex-stratified model. It is evident from Table IX.F-7 that the model was not significantly improved by adjusting for any of the other factors as a potential confounder: none produced a significantly better fit to the data. Since the estimated slope was virtually unaffected by such adjustments, it does not appear that omitting these factors introduces any important bias in the dose-response results.

Table IX.F-7. Confounding and Effect Modification by Sex, Age at Exposure or HTDS Examination, Estimated Thyroid Dose from NTS, History of Any Cancer Other Than Thyroid and Interview Type: Any Thyroid Nodule

		Estimated Dose-Response Coefficient (per Gy)				
Covariate (0=No, 1=Yes)	Yes / Total	Unadjusted	Adjusted for Confounding	Including Effect Modification Group 0	Including Effect Modification Group 1	P
Female?	1622 / 3191	−.087 ± .301 (−.808, .634)	Not Applicable	−.311 ± .555 (−1.70, 1.08)	.016 ± .357 (−.875, .908)	.62
Prenatal exposure?	1034 / 3191	−.087 ± .301 (−.808, .634)	−.150 ± .308 (−.943, .643)	−.064 ± .340 (−.962, .834)	−.461 ± .685 (−2.26, 1.35)	.58
1st exposure before age 180 days?	1478 / 3191	−.087 ± .301 (−.808, .634)	−.103 ± .310 (−.901, .695)	−.292 ± .510 (−1.64, 1.05)	.008 ± .375 (−.983, .998)	.64
Age at exam > 50?	2001 / 3191	−.087 ± .301 (−.808, .634)	−.180 ± .313 (−.987, .627)	−.064 ± .570 (−1.57, 1.44)	−.227 ± .373 (−1.21, .757)	.81
NTS ^{131}I dose > 5.3 mGy?	1566 / 3187	−.092 ± .303 (−.816, .633)	−.084 ± .308 (−.878, .711)	.183 ± .375 (−.805, 1.17)	−.551 ± .567 (−2.05, .944)	.26
History of any cancer other than thyroid?	248 / 3186	−.087 ± .301 (−.807, .634)	−.083 ± .303 (−.863, .698)	−.234 ± .344 (−1.14, .673)	.483 ± .560 (−.994, 1.96)	.31
Expanded in-person interview?	1212 / 3191	−.087 ± .301 (−.808, .634)	.012 ± .303 (−.768, .793)	.050 ± .459 (−1.16, 1.26)	−.016 ± .408 (−1.09, 1.06)	.91

Entries in the table are estimate ± standard error for the regression coefficient, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

Tables IX.F-8 and IX.F-9 display similar results from analyses including history of medical or dental x-ray exposure or occupational exposure as potential confounding or effect modifying factors. The estimates of the regression coefficient calculated with adjustment for confounding are all close to the unadjusted estimates. Moreover the adjusted estimates all remained less than zero. Thus there was no evidence that a confounding effect of any of these covariates has obscured a positive dose-response for the outcome of any thyroid nodule.

There is no evidence of any statistically significant effect modification by any of the covariates in Tables IX.F-8 and IX.F-9, with two possible exceptions.

- The estimated dose-response coefficient was markedly negative (−2.22) for the 398 participants with histories of intravenous pyelograms (IVPs), but not markedly different for the majority of participants without such histories (0.113 with confidence interval ranging from −0.675 to 0.900; p = 0.040).

- The estimated dose-response coefficient was markedly negative (−2.75) for the 442 participants with histories of occupations that might have involved exposure to radioactive materials or x-rays, but not markedly different for the majority of participants without such histories (0.103 with confidence interval ranging from −0.711 to 0.918; p= 0.031).

The statistical significance of these differences must be interpreted with caution due to the large number of such comparisons that were performed. Moreover, neither of these two covariates identified a group of participants with a significantly positive dose-response.

Table IX.F-8. Confounding and Effect Modification by History of Medical and Dental Radiation Exposures: Any Thyroid Nodule

Have You Ever Had: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy) Unadjusted	Adjusted for Confounding	Including Effect Modification Group 0	Including Effect Modification Group 1	P
CAT scan of the upper body?	775 / 3149	−.115 ± .305 (−.844, .615)	−.110 ± .303 (−.890, .669)	.073 ± .306 (−.735, .881)	−1.32± .89 (−3.66, 1.02)	.12
Diagnostic x-rays of the head?	1191 / 3155	−.060 ± .300 (−.777, .658)	−.062 ± .299 (−.833, .709)	.131 ± .339 (−.763, 1.02)	−.554 ± .602 (−2.14, 1.03)	.31
Diagnostic x-rays of the neck?	966 / 3167	−.081 ± .301 (−.801, .640)	−.104 ± .302 (−.882, .675)	.111 ± .394 (−.928, 1.15)	−.393 ± .501 (−1.72,.930)	.42
Diagnostic x-rays of chest or upper body, including mammograms?	2821 / 3173	−.090 ± .302 (−.813, .632)	−.077 ± .302 (−.854, .700)	−.215 ± 1.23 (−3.46, 3.03)	−.068 ± .311 (−.888, .752)	.91
Diagnostic x-rays of the stomach or mid-back?	692 / 3120	−.109 ± .308 (−.848, .629)	−.111 ± .308 (−.906, .683)	−.138 ± .347 (−1.05, .778)	−.007 ± .675 (−1.79, 1.77)	.86
Barium enema?	825 / 3159	−.094 ± .301 (−.815, .628)	−.093 ± .302 (−.869, .684)	−.056 ± .348 (−.974, .862)	−.198 ± .604 (−1.79, 1.40)	.84
Upper GI?	1146 / 3177	−.110 ± .304 (−.838, .618)	−.113 ± .304 (−.897, .671)	.041 ± .364 (−.920, 1.00)	−.401 ± .535 (−1.81, 1.01)	.49
Intravenous pyelogram?	398 / 3157	−.090 ± .303 (−.814, .635)	−.078 ± .303 (−.859, .703)	.113 ± .299 (−.675, .900)	−2.22 ± 1.22 (−5.45, 1.01)	.040
Fluoroscopy of the upper body?	246 / 3161	−.057 ± .300 (−.776, .662)	−.061 ± .301 (−.836, .714)	.054 ± .300 (−.738, .847)	−1.90 ± 1.45 (−5.73, 1.92)	.14
Nuclear scan (excluding thyroid scan)?	217 / 3162	−.086 ± .302 (−.809, .636)	−.084 ± .302 (−.862, .694)	.021 ± .300 (−.771, .814)	−2.30 ± 1.71 (−6.80, 2.20)	.13
Dental x-rays that did not usually include a lead shield over the neck area?	1648 / 3191	−.087 ± .301 (−.808, .634)	−.088 ± .301 (−.865, .688)	.344 ± .375 (−.645, 1.33)	−.662 ± .506 (−2.00, .674)	.103

Entries in the table are estimate ± standard error for the regression coefficient, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

Table IX.F-9. Confounding and Effect Modification by Occupational History: Any Thyroid Nodule

Have You Ever Worked in Any of the Following: (0=No, 1=Yes)	Yes/ Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted	Adjusted for Confounding	Including Effect Modification		P
				Group 0	Group 1	
Any metal industry?	238 / 3191	–.087 ± .301 (–.808, .634)	–.079 ± .300 (–.853, .695)	–.085 ± .305 (–.890, .720)	.144 ± 1.77 (–4.53, 4.82)	.90
Any nuclear facility?	371 / 3191	–.087 ± .301 (–.808, .634)	–.080 ± .303 (–.860, .701)	–.187 ± .341 (–1.09, .713)	.418 ± .661 (–1.33, 2.16)	.44
Any other industry or occupation where you may have been exposed to radioactive materials or x-rays?	442 / 3191	–.087 ± .301 (–.808, .634)	–.105 ± .304 (–.887, .677)	.103 ± .309 (–.711, .918)	–2.75 ± 1.53 (–6.80, 1.30)	.031
Any of the above industries or occupations?	892 / 3191	–.087 ± .301 (–.808, .634)	–.057 ± .301 (–.832, .717)	–.118 ± .358 (–1.06, .827)	–.097 ± .547 (–1.35, 1.54)	.74

Entries in the table are estimate ± standard error for the regression coefficient, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

Table IX.F-10 displays the results of analyses of possible confounding or effect modification by smoking variables. There was no evidence that the dose-response was significantly confounded by either smoking variable, or that there was a dose-response that differed significantly according to smoking history.

Table IX.F-10. Confounding and Effect Modification by Smoking: Any Thyroid Nodule

Have You Ever Smoked Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted	Adjusted for Confounding	Including Effect Modification		P
				Group 0	Group 1	
Cigarettes (unfiltered or filtered)?	1854 / 3183	–.083 ± .301 (–.803, .637)	–.083 ± .301 (–.858, .691)	–.198 ± .520 (–1.57, 1.17)	–.024 ± .364 (–.984, .936)	.78
Any of cigarettes, cigar or pipe?	1900 / 3183	–.083 ± .301 (–.803, .637)	–.083 ± .301 (–.858, .691)	–.105 ± .523 (–1.48, 1.27)	–.072 ± .367 (–1.04, .895)	.96

Entries in the table are estimate ± standard error for the regression coefficient, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

F.2.k. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for any thyroid nodule are shown in Figure IX.F-1 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in that figure are calculated at the 98.33% confidence level, i.e., are adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates.

While the point estimate of the slope is greater than 0 for 32 of the 100 realizations, the confidence interval includes 0 for all 100 of the realizations. Also shown in Figure IX.F-1 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean and mean of each participant's 100 dose realizations. In summary, for none of the 100 realizations of the estimated doses was there a statistically significant dose-response, and for the majority of realizations the estimated slope was less than 0.

Figure IX.F-1. Plot of Estimated Slope and 95% CI by Dose Realization: Any Thyroid Nodule

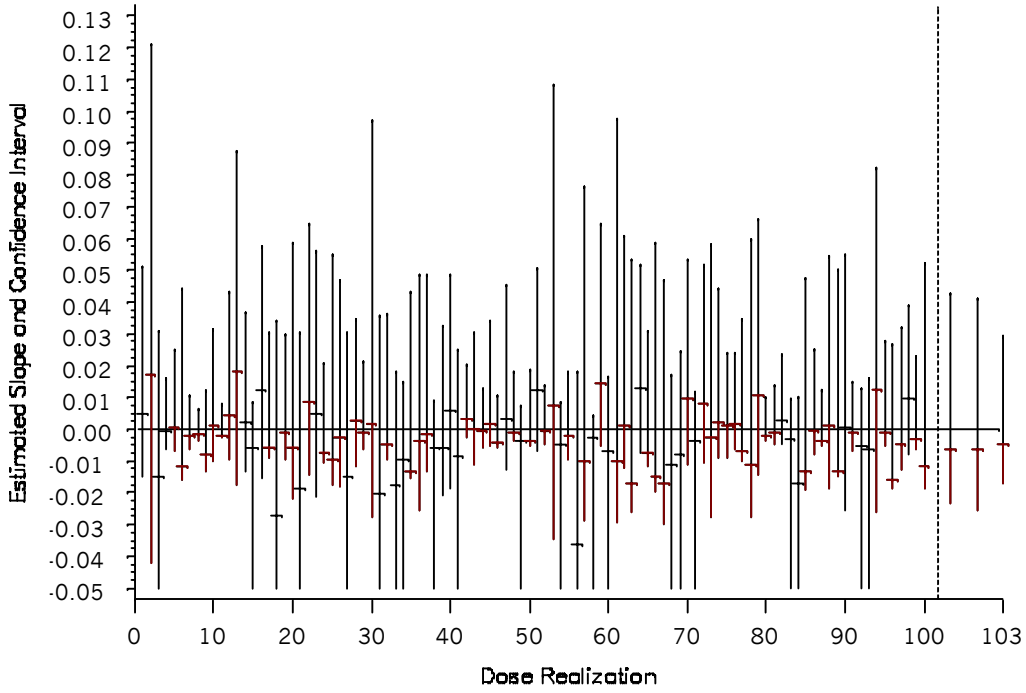
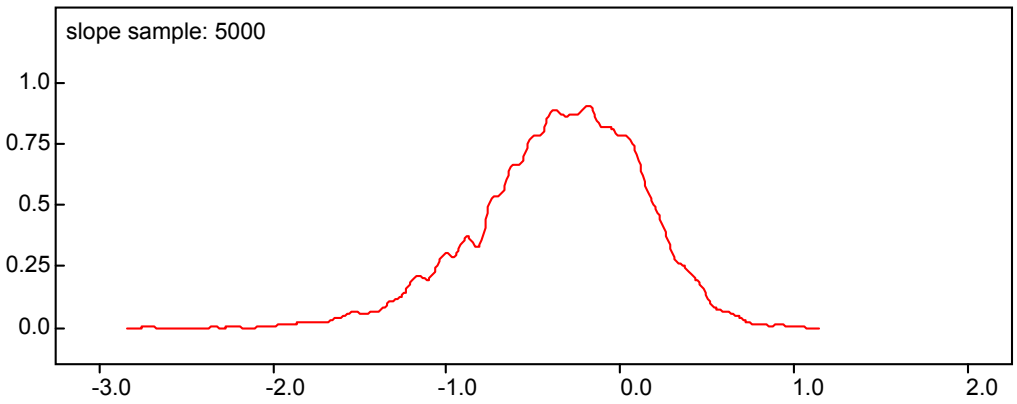


Figure IX.F-2 displays the distribution of the 5000 estimates of the logistic regression coefficient obtained by the simulation procedures described in section VIII.C.2.b.3 above. It is evident from the figure that most of the estimates were between about –2.0 and 1.0. The estimate was less than or equal to 0 for 3800 of the 5000 replications, implying an empirical one-tailed p-value of 0.76. The median estimate was –.303 and the upper and lower percentiles corresponding to the Bonferroni-adjusted 95% confidence interval were –1.65 and 0.62. These may be compared to the estimates of –.09 with confidence interval (–.81, .63) obtained using the median dose estimates without adjustment for uncertainty. Thus, this method of adjusting the estimated logistic regression coefficient for the uncertainty in the dose estimates did not provide evidence that the cumulative incidence of any thyroid nodule increased with increasing dose.

Figure IX.F-2. Distribution of Simulation Estimates of Logistic Regression Coefficient: Any Thyroid Nodule



G. Hypothyroidism

G.1. Occurrence of Hypothyroidism

The primary and alternative definitions for hypothyroidism were as follows:

- Primary definition: HTDS evaluation or medical records with supporting documentation (267 cases)
- Alternative definition #1: HTDS evaluation or medical records with or without supporting documentation (372 cases)
- Alternative definition #2: HTDS evaluation, any medical records, or inferred from past/current therapy (402 cases)
- Alternative definition #3: Any diagnosis or participant/respondent report (595 cases).

Two hundred and sixty-seven (7.8%) living evaluable participants had a diagnosis of hypothyroidism based on the HTDS evaluation or on medical records with supporting documentation, with 204 (11.7%) women and 63 (3.7%) men having this diagnosis, respectively (Table IX.G-1). An additional 105 (3.1%) living evaluable participants had a diagnosis of hypothyroidism based on medical records but without supporting documentation, and 30 (0.9%) were inferred from past or current therapy. There were 193 (5.6%) reports of hypothyroidism from the participant or his/her CATI respondent.

It should be noted that Alternative definition #1 includes cases from medical records *without* supporting documentation; this category includes many participants who have been treated with thyroid hormone for many years, had normal thyroid function on the HTDS lab evaluation, and yet had no early documentation of an elevated TSH in their medical records. This category therefore very likely includes an unknown number of valid diagnoses for hypothyroidism for which adequate diagnostic information was not available.

Table IX.G-1. Basis for Diagnosis of Hypothyroidism, by Sex

	Female		Male		Total	
Diagnosis of Hypothyroidism	No.	%	No.	%	No.	%
Yes	481	27.5	114	6.7	595	17.3
▪ HTDS evaluation	97	5.6	49	2.9	146	4.2
▪ Medical records *with* supporting documentation	107	6.1	14	0.8	121	3.5
▪ Medical records *without* supporting documentation	91	5.2	14	0.8	105	3.1
▪ Inferred from past/current therapy	27	1.5	3	0.2	30	0.9
▪ Participant/respondent report	159	9.1	34	2.0	193	5.6
No	1250	71.6	1575	93.0	2825	82.1
Unknown	16	0.9	4	0.2	20	0.6
Total	1747	100.0	1693	100.0	3440	100.0

Twenty living evaluable participants were classified as "unknown" with regard to diagnosis of hypothyroidism. These twenty did not have medical record reports of such a diagnosis. Seven of these 20 participants did not have a blood draw, thirteen had participant/respondent report of unknown thyroid disease with 11 taking some kind of medication for this unknown thyroid disease. Therefore, no HTDS evaluation could be made for these twenty participants who were included as non-cases in analyses of the dose-response for hypothyroidism.

Of those with a diagnosis of hypothyroidism, 531 (89.2%) had no known contributing cause (Table IX.G-2). However, among those with a contributing cause, about half were due to thyroid or parathyroid surgery, followed by ^{131}I therapy (21), and lithium therapy (6). Ten participants had some

other contributing cause, with four of the 10 being unknown or uncertain, while two were related to subacute thyroiditis.

Table IX.G-2. Frequency Distribution of Possible Contributing Causes of Hypothyroidism, by Sex

	Female		Male		Total	
Contributing Cause	No.	%	No.	%	No.	%
No Known Contributing Cause	427	88.8	104	91.2	531	89.2
Contributing Cause	54	11.2	10	8.8	64	10.8
^{131}I therapy	20	4.2	1	0.9	21	3.5
Thyroid/parathyroid surgery	25	5.2	5	4.4	30	5.0
Lithium Therapy	4	0.8	2	1.8	6	1.0
Other	8	1.7	2	1.8	10	1.7
Total	481	100.0	114	100.0	595	100.0

Note: A participant can have more than one possible contributing cause

G.1.a. Permanent Hypothyroidism

An additional outcome of hypothyroidism was defined to exclude those with transient hypothyroidism. Transient (temporary) hypothyroidism can occur from certain types of thyroiditis such as viral subacute thyroiditis or postpartum thyroiditis. Transient forms of hypothyroidism usually resolve completely and do not require further treatment. In contrast, permanent hypothyroidism, such as that produced from Hashimoto's thyroiditis, ^{131}I therapy, or thyroid surgery, requires lifelong thyroid hormone replacement. The definition of permanent hypothyroidism included participants with a diagnosis of hypothyroidism based on the HTDS evaluation. Permanent hypothyroidism also included those based on medical records with supporting documentation, excluding those who had a normal TSH value at the time of the HTDS clinic and were not currently on thyroid hormone replacement. Two hundred and fifty seven participants (7.5%) had a diagnosis of permanent hypothyroidism (Table IX.G-3). These cases represented 96% of the cases of hypothyroidism according to the primary definition (i.e., diagnosed from the HTDS evaluation or medical records with supporting documentation).

Table IX.G-3. Permanent Hypothyroidism, by Sex

	Female		Male		Total	
Permanent Hypothyroidism	No.	%	No.	%	No.	%
Yes	196	11.2	61	3.6	257	7.5
No	1551	88.8	1632	96.4	3183	92.6
Total	1747	100.0	1693	100.0	3440	100.0

G.2. Analysis of Hypothyroidism Risk

G.2.a. Primary Analysis

Of the 267 participants with a diagnosis of hypothyroidism based on the HTDS examination or medical records with supporting documentation, 21 were out-of-area participants. The number of cases and proportion with hypothyroidism are shown by sex, in-area status, and dose group in Tables IX.G-4 and IX.G-5.

Table IX.G-4. Diagnoses of Hypothyroidism by Sex, Dose Category, and Basis for Diagnosis

A. Female

Thyroid Radiation Dose	Living Evaluable Female	Primary Definition: Cases Based on HTDS or Medical Record with Supporting Document		1st Alternative Definition: Cases Based on HTDS or Med. Rec. with or without Supporting Document		2nd Alternative Definition: Cases Based on HTDS or Medical Record, or Inferred from Medication		3rd Alternative Definition: Cases Based on Any Diagnosis or Participant or CATI Report	
(mGy)	No.	No.	%	No.	%.	No.	%	No.	%
OOA	125	14	11.2	20	125	23	18.4	31	24.8
< 10	182	19	10.4	26	182	29	15.9	42	23.1
10-49	320	34	10.6	42	320	47	14.7	67	20.9
50-99	313	40	12.8	64	313	68	21.7	106	33.9
100-149	220	22	10.0	34	220	37	16.8	61	27.7
150-199	126	14	11.1	21	126	21	16.7	31	24.6
200-299	139	20	14.4	28	139	32	23.0	46	33.1
300-399	144	22	15.3	31	144	34	23.6	49	34.0
400-999	171	18	10.5	27	171	29	17.0	46	26.9
1000+	7	1	14.3	2	7	2	28.6	2	28.6
Total	1747	204	11.7	295	1747	322	18.4	481	27.5

OOA = out of area participant

B. Male

Thyroid Radiation Dose	Living Evaluable Male	Primary Definition: Cases Based on HTDS or Medical Record with Supporting Document		1st Alternative Definition: Cases Based on HTDS or Med. Rec. with or without Supporting Document		2nd Alternative Definition: Cases Based on HTDS or Medical Record, or Inferred from Medication		3rd Alternative Definition: Cases Based on Any Diagnosis or Participant or CATI Report	
(mGy)	No.	No.	%	No.	%	No.	%	No.	%
OOA	124	7	5.6	7	125	31	125	10	8.1
< 10	186	8	4.3	10	182	42	182	13	7.0
10-49	314	10	3.2	12	320	67	320	18	5.7
50-99	310	11	3.5	12	313	106	313	19	6.1
100-149	171	7	4.1	8	220	61	220	12	7.0
150-199	109	2	1.8	2	126	31	126	5	4.6
200-299	148	7	4.7	8	139	46	139	11	7.4
300-399	160	7	4.4	10	144	49	144	13	8.1
400-999	154	4	2.6	7	171	46	171	11	7.1
1000+	17	0	--	1	7	2	7	2	11.8
Total	1693	63	3.7	77	1747	481	1747	114	6.7

OOA = out of area participant

Table IX.G-5. Additional Disease Outcomes Related to Hypothyroidism by Sex and Estimated Dose (cases based on HTDS examination or medical records with supporting documentation only)

A. Female

Thyroid Radiation Dose	Living Evaluable Female	Permanent Hypothyroidism	
(mGy)	No.	No.	%
OOA	125	14	11.2
< 10	182	18	9.9
10-49	320	34	10.6
50-99	313	36	11.5
100-149	220	20	9.1
150-199	126	14	11.1
200-299	139	19	13.7
300-399	144	22	15.3
400-999	171	18	10.5
1000+	7	1	14.3
Total	1747	196	11.2

B. Male

Thyroid Radiation Dose	Living Evaluable Male	Permanent Hypothyroidism	
(mGy)	No.	No.	%
OOA	124	7	5.6
< 10	186	7	3.8
10-49	314	10	3.2
50-99	310	11	3.5
100-149	171	6	3.5
150-199	109	2	1.8
200-299	148	7	4.7
300-399	160	7	4.4
400-999	154	4	2.6
1000+	17	0	--
Total	1693	61	3.6

OOA = out of area participant

Parameter estimates for the linear dose-response model based on the 3191 in-area participants are shown in row 1 of Table IX.G-6 below. Based on maximum likelihood analysis of the sex-stratified linear probability model, the estimated slope B was slightly less than zero (–0.006 per Gy) with Bonferroni-adjusted 95% CI ranging from less than –0.016 to 0.047 per Gy, providing no evidence that cumulative incidence increased with increasing dose (one-tailed p = 0.61). The corresponding estimated background rates for diagnosis of benign thyroid nodule were 0.118 with confidence interval (0.097, 0.139) for women and 0.037 with confidence interval (0.023, 0.050) for men. Similar results were obtained when the model was fit by the method of least squares using ungrouped or grouped data, although the estimates of the slope were slightly greater than zero: 0.006 per Gy with confidence interval (–0.044, 0.056 per Gy) with ungrouped data, and 0.002 per Gy with confidence interval (–0.055, 0.060 per Gy) with grouped data (Table IX.G-6, rows 2 and 3 respectively).

Table IX.G-6. Dose-Response Results for Diagnoses of Hypothyroidism

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
1.	Primary definition (HTDS evaluation or medical record with documentation)	Linear	Primary	None	MLE	.118 ± .009 (.097, .139)	.037 ± .006 (.023, .050)	–.006 ± .019 (< –.016, .047)	0.61
2.	Primary definition	Linear	Primary	None	LSU	.116 ± .008 (.098, .134)	.035 ± .008 (.016, .053)	.006 ± .021 (–.044, .056)	0.39
3.	Primary definition	Linear	Primary	None	LSG	.117 ± .008 (.098, .135)	.035 ± .008 (.016, .054)	.002 ± .024 (–.055, .060)	0.46
4.	Alternative def. #1 (HTDS evaluation or medical records with or without documentation)	Linear	Primary	None	MLE	.165 ± .010 (.141, .189)	.040 ± .006 (.025, .055)	.026 ± .023 (< -.020, .086)	0.12
5.	Alternative def. #2 (HTDS evaluation or medical record, or inferred from medication	Linear	Primary	None	MLE	.180 ± .010 (.155, .205)	.042 ± .006 (.026, .057)	.025 ± .024 (< -.020, .087)	0.13

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued next page

Table IX.G-6. Dose-Response Results for Diagnoses of Hypothyroidism (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
6.	Alternative def. #3 (Any diagnosis or participant/respondent report)	Linear	Primary	None	MLE	.271 ± .012 (.242, .300)	.060 ± .008 (.042, .078)	.038 ± .028 (-.023, .108)	0.076
7.	Permanent hypothyroidism	Linear	Primary	None	MLE	.112 ± .009 (.092, .133)	.035 ± .006 (.021, .048)	-.001 ± .020 (< -.015, .053)	0.52
8.	Primary definition	LQ	Primary	None	LSU	.117 ± .008 (.097, .138)	.036 ± .008 (.015, .057)	Lin: –.006 ± .035 (–.094, .082) Quad: .010 ± .023 (–.049, .068)	Quad: 0.68
9.	Primary definition	Logistic	Primary	None	MLE	.116 (.095, .140)	.035 (.025, .049)	.08 ± .29 (-.62, .78)	0.39

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.G-6. Dose-Response Results for Diagnoses of Hypothyroidism (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
10.	Alternative def. #1	Logistic	Primary	None	MLE	.161 (.137, .188)	.042 (.031, .056)	.37 ± .23 (-.19, .93)	0.065
11.	Alternative def. #2	Logistic	Primary	None	MLE	.176 (.151, .204)	.043 (.032, .058)	.34 ± .23 (-.22, .89)	0.08
12.	Alternative def. #3	Logistic	Primary	None	MLE	.266 (.236, .298)	.062 (.049, .080)	.33 ± .21 (-.16, .83)	0.055
13.	Primary definition	Linear	Primary	Exclude dose > 1000 mGy	MLE	.117 ± .009 (.096, .139)	.036 ± .006 (.022, .051)	-.002 ± .023 (< -.047, .060)	0.53
14.	Primary definition	Linear	Primary	Exclude dose > 400 mGy	MLE	.113± .010 (.089,.136)	.032 ± .006 (.017, .047)	.047 ± .041 (-.045, .151)	0.12

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.G-6. Dose-Response Results for Diagnoses of Hypothyroidism (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
15.	Primary definition	Linear	Primary	Exclude Ok and F/S geostrata	MLE	.115 ± .009 (.093, .138)	.031 ± .006 (.017, .045)	.004 ± .021 (<-.014, .060)	0.42
16.	Primary definition	Linear	Alt. #1	None	MLE	.120 ± .009 (.098, .141)	.038 ± .006 (.024, .051)	-.011 ± .017 (< -.016, .037)	0.74
17.	Primary definition	Linear	Alt. #2	None	MLE	.117 ± .009 (.096, .138)	.036 ± .006 (.022, .050)	.0002 ± .020 (< -.017, .053)	0.50
18.	Primary definition	Linear	Primary	Include OOA (scoping analysis #1)	MLE	.118 ± .008 (.098, .138)	.038 ± .005 (.025, .051)	-.008 ± .019 (< -.016, .044)	0.66
19.	Primary definition	Linear	Primary	Include OOA (scoping analysis #2)	MLE	.118 ± .008 (.098, .138)	.039 ± .005 (.026, .052)	-.010 ± .018 (< -.016, >.041)	0.69

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

G.2.b. Alternative Definitions for Diagnosis of Hypothyroidism

Each of the three alternative definitions of hypothyroidism (section IX.G.1. above) added substantial numbers of cases.

Parameter estimates for the linear dose-response model are shown in rows 4 and 5 of Table IX.G-6 above for each of the three alternative definitions of hypothyroidism. For none of the three alternative definitions was there a clearly statistically significant increase of cumulative incidence with increasing dose.

G.2.c. Permanent Hypothyroidism

In the analyses described above, participants with transient hypothyroidism were included among the cases. An additional analysis was performed in which participants with transient hypothyroidism only were excluded from the cases. The results are shown in row 7 of Table IX.G-6 above. The cumulative incidence of permanent hypothyroidism decreased slightly with increasing dose, with an estimated slope of –0.001 per Gy and confidence interval ranging from less than –0.015 to 0.053 per Gy ($p = 0.52$).

G.2.d. Alternative Dose-Response Functions

As shown in row 8 of Table IX.G-6, the estimated regression coefficient for the dose-squared term in the linear-quadratic dose-response model [5] was 0.010 with Bonferroni-adjusted 95% confidence interval ranging from –0.049 to 0.068. Thus the addition of a quadratic term did not significantly improve the fit of the model ($p = 0.68$).

In the analysis of hypothyroidism based on the HTDS or medical records with supporting documentation, i.e., the primary criterion for defining cases with hypothyroidism, the regression parameter for the effect of dose in the sex-stratified logistic regression model [2] was estimated as 0.08, with Bonferroni-adjusted 95% confidence limits ranging from –0.62 to 0.78 (Table IX.G-6, row 9). Thus in the primary analysis of hypothyroidism, there was no evidence from the logistic regression model that cumulative incidence increased significantly with increasing dose ($p = 0.39$). However in logistic regression analyses using the alternative criteria for defining cases with hypothyroidism, the estimated regression coefficients were larger but did not achieve statistical significance (see rows 10, 11 and 12 of Table IX.G-6).

G.2.e. Effect of Excluding Participants in High Dose Categories

When those with an estimated dose > 1000 mGy were excluded, the estimated slope B was –0.002 per Gy with Bonferroni-adjusted 95% confidence interval ranging from less than –0.047 to 0.060 per Gy, providing no evidence that the cumulative incidence of hypothyroidism increased with increasing dose ($p = 0.53$; Table IX.G-6, row 13). When participants with estimated dose > 400 mGy were excluded, the estimated slope was 0.047 per Gy with confidence interval ranging from –0.045 to 0.151 per Gy, and there was no evidence that the cumulative incidence increased with increasing dose ($p = 0.12$; Table IX.G-6, row 14).

G.2.f. Effect of Excluding Okanogan and Ferry/Stevens Geostrata

When Okanogan and Ferry/Stevens geostrata were excluded from the analyses, the estimated slope of the dose-response increased from –0.006 per Gy to the slightly positive value of 0.004 per Gy,

with confidence interval ranging from less than –0.014 to 0.060 per Gy (Table IX.G-6, row 15). Thus the cumulative incidence of hypothyroidism did not increase significantly with increasing dose (p = 0.42).

G.2.g. Analysis of Hypothyroidism in Relation to Alternative Dose Estimates

In the analysis using the first set of alternative dose estimates (Table IX.G-6, row 16), the estimated slope B was –0.011 per Gy, with Bonferroni-adjusted 95% confidence interval ranging from less than -0.016 to 0.037 per Gy, providing no evidence that the cumulative incidence of hypothyroidism increased with increasing dose (p = 0.74). The results for the second set of alternative dose estimates were similar (Table IX.G-6, row 17), with an estimated slope B of 0.0002 per Gy, with Bonferroni-adjusted 95% confidence interval ranging from less than -0.017 to 0.053 per Gy, providing no evidence that the cumulative incidence increased with increasing dose (p = 0.50).

G.2.h. Scoping Analysis Regarding Out-of-Area Participants

See section VIII.C.1.a.3 for a description of the scoping analyses that were performed to assess the possible impact of including the 249 out-of-area participants. As summarized in rows 18 and 19 of Table IX.G-6, in both analyses the inclusion of the out-of-area participants slightly decreased the estimated slope of the dose-response, but did not materially change the dose-response results. In particular, the estimated slope of the dose-response was slightly negative in both scoping analyses, providing no evidence that cumulative incidence increased with increasing dose.

G.2.i. Analysis of Hypothyroidism in Relation to Alternative Representations of Exposure

In the analyses by geostratum and by dichotomous exposure, the sex and age-adjusted comparisons of cumulative incidence were performed as described in section VIII.C.2.a.2.

G.2.i.1. Analysis by Geostratum

The proportions of women with hypothyroidism (HTDS or medical record with documentation) ranged from 11/75 (14.7%) in the Okanogan geostratum to 15/165 (9.1%) in the Adams geostratum (Table IX.G-7). For men they ranged from 7/70 (10.0%) in the Ferry/Stevens geostratum to 8/501 (1.6%) in the Pasco/Kennewick geostratum. However the heterogeneity among the nine geostrata was not statistically significant (p = 0.51). Hypothyroidism was somewhat more common in the Okanogan and Ferry/Stevens geostrata (13.3% and 8.2% for women and men, respectively) compared to the other geostrata (11.5% and 3.3%, respectively), but this difference was not statistically significant (p = 0.12).

Table IX.G-7. Diagnoses of Hypothyroidism Based on the HTDS Evaluation or on Medical Records with Supporting Documentation

	Female			Male			Total		
Geostratum	No.	Cases	%	No.	Cases	%	No.	Cases	%
Richland	179	21	11.7	173	7	4.0	352	28	8.0
Pasco/Kennewick	508	58	11.4	501	8	1.6	1009	66	6.5
Benton County	376	45	12.0	358	16	4.5	734	61	8.3
Franklin County	73	10	13.7	76	6	7.9	149	16	10.7
Adams County	165	15	9.1	156	5	3.2	321	20	6.2
Walla Walla (city)	133	16	12.0	131	3	2.3	264	19	7.2
Walla Walla County	170	20	11.8	164	7	4.3	334	27	8.1
Okanogan County	75	11	14.7	64	4	6.3	139	15	10.8
Ferry/Stevens Counties	68	8	11.8	70	7	10.0	138	15	10.9
Total	1747	204	11.7	1693	63	3.7	3440	267	7.8

Because of the large numbers of cases added by the alternative criteria for defining cases of hypothyroidism (see IX.G-1 above), results for the alternative definitions of hypothyroidism are also presented (Tables IX.G-8 to IX.G-10, below). Generally, similar degrees of heterogeneity among the geostrata were observed in the analyses using the alternative definitions as compared to the primary definition, ($p = 0.57$, $p = 0.55$, and $p = 0.017$ for the first, second and third alternative definitions, respectively). Only when diagnoses that were reported by participants or CATI respondents but not confirmed by the HTDS evaluation were included (Table IX.G-10) was there evidence of significant heterogeneity among geostrata. The tendency toward higher proportions of cases in the Okanogan and Ferry/Stevens geostrata was observed in all three alternatives.

Table IX.G-8. Diagnoses of Hypothyroidism Based on the HTDS Evaluation or Medical Records with or without Supporting Documentation (1st Alternative Definition), by Geostratum and Sex

	Female			Male			Total		
Geostratum	No.	Cases	%	No.	Cases	%	No.	Cases	%
Richland	179	28	15.6	173	7	4.0	352	35	9.9
Pasco/Kennewick	508	87	17.1	501	15	3.0	1009	102	10.1
Benton County	376	60	16.0	358	16	4.5	734	76	10.4
Franklin County	73	13	17.8	76	6	7.9	149	19	12.8
Adams County	165	27	16.4	156	6	3.8	321	33	10.3
Walla Walla (city)	133	20	15.0	131	5	3.8	264	25	9.5
Walla Walla County	170	32	18.8	164	9	5.5	334	41	12.3
Okanogan County	75	17	22.7	64	5	7.8	139	22	15.8
Ferry/Stevens Counties	68	11	16.2	70	8	11.4	138	19	13.8
Total	1747	295	16.9	1693	77	4.5	3440	372	10.8

Table IX.G-9. Diagnoses of Hypothyroidism Based on the HTDS Evaluation, or on Medical Records with or without Supporting Documentation (2nd Alternative Definition), or Inferred from Past/Current Therapy, by Geostratum and Sex

	Female			Male			Total		
Geostratum	No.	Cases	%	No.	Cases	%	No.	Cases	%
Richland	179	30	16.8	173	7	4.0	352	37	10.5
Pasco/Kennewick	508	95	18.7	501	16	3.2	1009	111	11.0
Benton County	376	65	17.3	358	18	5.0	734	83	11.3
Franklin County	73	14	19.2	76	6	7.9	149	20	13.4
Adams County	165	29	17.6	156	6	3.8	321	35	10.9
Walla Walla (city)	133	23	17.3	131	5	3.8	264	28	10.6
Walla Walla County	170	34	20.0	164	9	5.5	334	43	12.9
Okanogan County	75	19	25.3	64	5	7.8	139	24	17.3
Ferry/Stevens Counties	68	13	19.1	70	8	11.4	138	21	15.2
Total	1747	322	18.4	1693	80	4.7	3440	402	11.7

Table IX.G-10. Diagnoses of Hypothyroidism Based on Any Source (3rd Alternative Definition), by Geostratum and Sex

	Female			Male			Total		
Geostratum	No.	Cases	%	No.	Cases	%	No.	Cases	%
Richland	179	47	26.3	173	10	5.8	352	57	16.2
Pasco/Kennewick	508	135	26.6	501	24	4.8	1009	159	15.8
Benton County	376	94	25.0	358	24	6.7	734	118	16.1
Franklin County	73	27	37.0	76	8	10.5	149	35	23.5
Adams County	165	41	24.8	156	8	5.1	321	49	15.3
Walla Walla (city)	133	39	29.3	131	8	6.1	264	47	17.8
Walla Walla County	170	55	32.4	164	16	9.8	334	71	21.3
Okanogan County	75	23	30.7	64	7	10.9	139	30	21.6
Ferry/Stevens Counties	68	20	29.4	70	9	12.9	138	29	21.0
Total	1747	481	27.5	1693	114	6.7	3440	595	17.3

G.2.i.2. Analysis by Dichotomous Exposure Variable

A total of 96 (7.6%) of the 1257 participants included in these analyses had a diagnosis of hypothyroidism based on the HTDS examination or a medical record with supporting documentation (see Table IX.G-11). These included 35/580 (6.0%) in the high exposure group and 61/677 (9.0%) in the low exposure group. Thus there was no evidence that the cumulative incidence of hypothyroidism was significantly higher in the high exposure group (p = 0.97).

Table IX.G-11. Diagnoses of Hypothyroidism Based on HTDS or Medical Record with Supporting Documentation, by Exposure Group and Sex

	Female			Male			Total		
Exposure Group	No	Cases	%	No.	Cases	%	No.	Cases	%
Low	351	40	11.4	326	21	6.4	677	61	9.0
High	298	31	10.4	282	4	1.4	580	35	6.0
Total	649	71	10.9	608	25	4.1	1257	96	7.6

Because of the large numbers of cases added by the three alternative definitions for hypothyroidism (see section IX.G.1 above), results for these alternatives are also presented briefly. The first alternative definition added 41 cases with diagnoses based on medical records without supporting documentation, for a total of 137 (10.9%). The second alternative added 11 cases with diagnoses inferred from medication, for a total of 148 (11.8%). The third alternative added 60 further cases with diagnoses reported by the participant or his/her CATI respondent, for a total of 208 (16.5%). In none of the alternative analyses was the cumulative incidence of hypothyroidism found to be elevated in the high exposure group ($p = 0.86$, 0.94, and 0.73 for the first, second, and third alternatives, respectively).

G.2.j. Confounding and Effect Modification

As described in section VIII above, additional sex-stratified logistic regression models were investigated to examine the possibility that the primary dose-response results might be influenced by confounding, and to search for factors that might modify a radiation dose-response. These analyses were based on the primary definition of hypothyroidism, i.e., those based on the HTDS evaluation or on medical records with supporting documentation, and on the primary dose estimates. Table IX.G-12 displays results for models including sex, age at first exposure to Hanford ^{131}I (prenatal, or < 180 days), age at HTDS examination, estimated dose from the NTS, history of any cancer other than thyroid, and HTDS interview type. Note that sex was not analyzed as a possible confounder since its effect was already adjusted for in the sex-stratified model. It is evident from Table IX.G-12 that the model was not significantly improved by adjusting for any of the other factors as a potential confounder: none produced a significantly better fit to the data. Since the estimated slope was virtually unaffected by such adjustments, it does not appear that omitting these factors introduces any important bias in the dose-response results.

Table IX.G-12. Confounding and Effect Modification by Sex, Age at Exposure or HTDS Examination, Estimated Thyroid Dose from NTS, History of Any Cancer Other Than Thyroid, and Interview Type: Hypothyroidism

		Estimated Dose-Response Coefficient (per Gy)				
Covariate (0=No, 1=Yes)	Yes / Total	Unadjusted	Adjusted for Confounding	Including Effect Modification Group 0	Including Effect Modification Group 1	P
Female?	1622 / 3191	.082 ± .293 (–.620, .785)	Not Applicable	–.508 ± .704 (–2.27, 1.25)	.240 ± .321 (–.563, 1.04)	.31
Prenatal exposure?	1034 / 3191	.082 ± .293 (–.620, .785)	.063 ± .297 (–.702, .829)	.023 ± .346 (–.891, .936)	.187 ± .589 (–1.37, 1.74)	.81
1st exposure before age 180 days?	1478 / 3191	.082 ± .293 (–.620, .785)	.097 ± .297 (–.669, .863)	–.630 ± .568 (–2.14, .859)	.386 ± .317 (–.451, 1.22)	.11
Age at exam >50?	2001 / 3191	.082 ± .293 (–.620, .785)	.149 ± .292 (–.604, .902)	.369 ± .445 (–.805, 1.54)	.004 ± .386 (–1.01, 1.02)	.54
NTS ^{131}I dose > 5.3 mGy?	1563 / 3181	.084 ± .297 (–.627, .795)	.002 ± .311 (–.800, .804)	–.184 ± .422 (–1.30, .929)	.248 ± .445 (–.925, 1.42)	.49
History of any cancer other than thyroid?	248 / 3186	.083 ± .293 (–.620, .785)	.081 ± .293 (–.674, .836)	.043 ± .330 (–.828, .914)	.224 ± .614 (–1.40, 1.84)	.80
Expanded In-Person Interview?	1212 / 3191	.082 ± .293 (–.620, .785)	.119 ± .298 (–.648, .885)	–.603 ± .541 (–2.03, .826)	.456 ± .325 (–.402, 1.31)	.089

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

Tables IX.G-13 and IX.G-14 display similar results from analyses including history of medical or dental x-ray exposure or of occupational exposures as potential confounding or effect modifying factors. The estimates of the regression coefficient calculated with adjustment for confounding are all close to the unadjusted estimates. Thus there was no evidence that a confounding effect of any of these covariates has obscured a positive dose-response for hypothyroidism.

There is no evidence of any statistically significant effect modification by any of the covariates in Tables IX.G-13 and IX.G-14.

Table IX.G-13. Confounding and Effect Modification by History of Medical and Dental Radiation Exposures: Hypothyroidism

Have You Ever Had: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy) Unadjusted	Adjusted for Confounding	Including Effect Modification Group 0	Group 1	P
CAT scan of the upper body?	775 / 3149	.046 ± .297 (–.664, .757)	.034 ± .299 (–.736, .803)	–.044 ± .342 (–.945, .857)	.356 ± .675 (–1.42, 2.14)	.60
Diagnostic x-rays of the head?	1191 / 3155	.083 ± .296 (–.625, .791)	.080 ± .296 (–.684, .843)	.103 ± .362 (–.852, 1.06)	.033 ± .514 (–1.32, 1.39)	.91
Diagnostic x-rays of the neck?	966 / 3167	.059 ± .297 (–.652, .770)	.070 ± .296 (–.692, .832)	–.264 ± .437 (–1.42, .888)	.375 ± .376 (–.616, 1.37)	.27
Diagnostic x-rays of chest or upper body, including mammograms?	2821 / 3173	.080 ± .294 (–.624, .784)	.090 ± .294 (–.667, .847)	.907 ± 1.22 (–2.31, 4.12)	.047 ± .305 (–.757, .852)	.52
Diagnostic x-rays of the stomach or mid-back?	692 / 3120	.102 ± .294 (–.603, .807)	.099 ± .295 (–.660, .858)	–.087 ± .353 (–1.02, .843)	.678 ± .554 (–.785, 2.14)	.26
Barium enema?	825 / 3159	.085 ± .295 (–.622, .792)	.084 ± .295 (–.677, .845)	.272 ± .329 (–.595, 1.14)	–.495 ± .658 (–2.23, 1.24)	.28
Upper GI?	1146 / 3177	.066 ± .296 (–.642, .773)	.066 ± .295 (–.695, .827)	.146 ± .357 (–.796, 1.09)	–.090 ± .519 (–1.46, 1.28)	.71
Intravenous pyelogram?	398 / 3157	.007 ± .304 (–.721, .734)	.013 ± .304 (–.770, .796)	–.101 ± .337 (–.990, .788)	.804 ± .824 (–1.37, 2.98)	.32
Fluoroscopy of the upper body?	246 / 3161	.072 ± .295 (–.635, .779)	.078 ± .295 (–.682, .839)	.094 ± .304 (–.707, .895)	–.154 ± 1.19 (–3.28, 2.97)	.84
Nuclear scan (excluding thyroid scan)?	217 / 3162	.089 ± .292 (–.611, .789)	.098 ± .292 (–.653, .849)	.137 ± .293 (–.636, .909)	–1.19 ± 1.89 (–6.17, 3.80)	.46
Dental x-rays that did not usually include a lead shield over the neck area?	1648 / 3191	.082 ± .293 (–.620, .785)	.083 ± .294 (–.674, .839)	–.372 ± .493 (–1.67, .930)	.391 ± .357 (–.551, 1.33)	.20

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

Table IX.G-14. Confounding and Effect Modification by Occupational History: Hypothyroidism

Have You Ever Worked in Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted	Adjusted for Confounding	Including Effect Modification		P
				Group 0	Group 1	
Any metal industry?	238 / 3191	.082 ± .293 (–.620, .785)	.082 ± .293 (–.673, .838)	.154 ± .290 (–.611, .920)	–3.53 ± 2.83 (–11.0, 3.94)	.11
Any nuclear facility?	371 / 3191	.082 ± .293 (–.620, .785)	.071 ± .297 (–.695, .837)	.103 ± .315 (–.729, .935)	–.158 ± .871 (–2.45, 2.14)	.77
Any other industry or occupation where you may have been exposed to radioactive materials or x-rays?	442 / 3191	.082 ± .293 (–.620, .785)	.081 ± .294 (–.676, .838)	.164 ± .314 (–.665, .992)	–.476 ± .965 (–3.02, 2.07)	.49
Any of the above industries or occupations?	892 / 3191	.082 ± .293 (–.620, .785)	.079 ± .294 (–.678, .837)	.295 ± .326 (–.566, 1.16)	–.617 ± .686 (–2.43, 1.19)	.20

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

Table IX.G-15 displays the results of analyses of possible confounding or effect modification by smoking variables. There was no evidence that the dose-response was significantly confounded by either smoking variable, or that there was a dose-response that differed significantly according to smoking history.

Table IX.G-15. Confounding and Effect Modification by Smoking: Hypothyroidism

Have You Ever Smoked Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted	Adjusted for Confounding	Including Effect Modification		P
				Group 0	Group 1	
Cigarettes (unfiltered or filtered)?	1854 / 3183	.086 ± .294 (–.617, .790)	.099 ± .296 (–.663, .860)	.078 ± .474 (–1.17, 1.33)	.112 ± .377 (–.883, 1.11)	.96
Any of cigarettes, cigar or pipe?	1900 / 3183	.086 ± .294 (–.617, .790)	.098 ± .296 (–.663, .860)	.071 ± .476 (–1.18, 1.33)	.116 ± .376 (–.875, 1.11)	.94

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

G.2.k. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for hypothyroidism are shown in Figure IX.G-1 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in that figure are calculated at the 98.33% confidence level, i.e., are adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates. While the point estimate of the slope is greater than 0 for 32 of the 100 realizations, the confidence interval

includes 0 for all 100 realizations. Also shown in Figure IX.G-1 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean and mean of each participant's 100 dose realizations. In summary, for none of the 100 realizations of the estimated doses was there a statistically significant dose-response, and for the majority of realizations the estimated slope was less than 0.

Figure IX.G-1 Plot of Estimated Slope and 95% CI by Dose Realization: Hypothyroidism

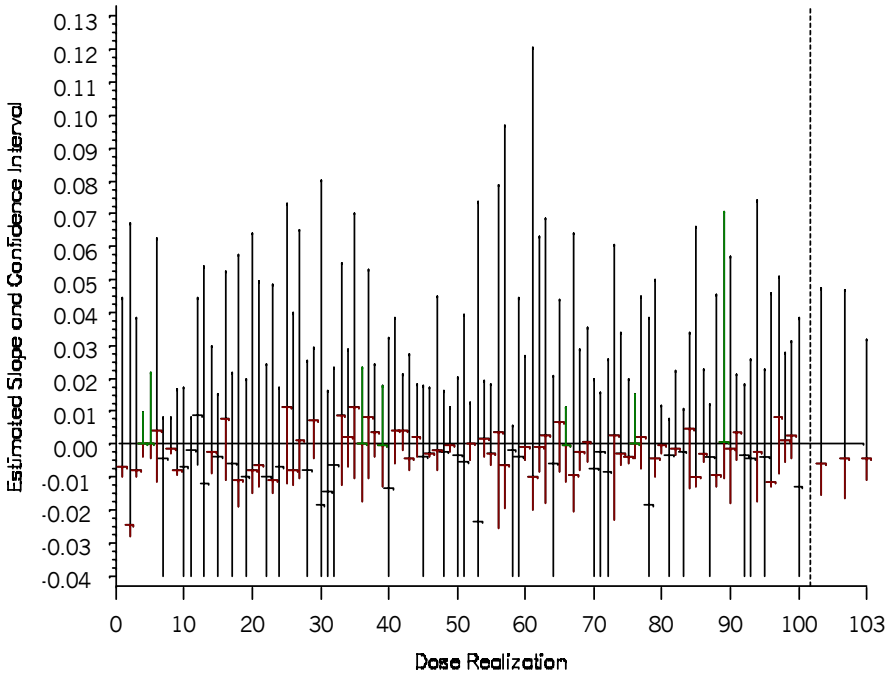
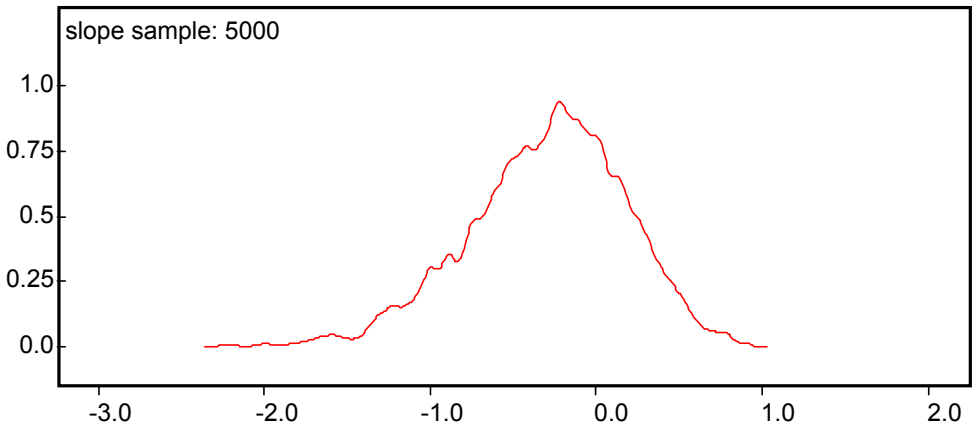


Figure IX.G-2 displays the distribution of the 5000 estimates of the logistic regression coefficient obtained by the simulation procedures described in section VIII.C.2.b.3 above. It is evident from the figure that most of the estimates were between about –2.0 and 1.0. The estimate was less than or equal to 0 for 2368 of the 5000 replications, implying an empirical one-tailed p-value of 0.47. The median estimate was 0.028, and the upper and lower percentiles corresponding to the Bonferroni-adjusted 95% confidence interval were –1.40 and 0.86. These may be compared to the estimates of –.08 with confidence interval (–0.62, 0.78) obtained using the median dose estimates without adjustment for uncertainty. Thus, this method of adjusting the estimated logistic regression coefficient for the uncertainty in the dose estimates did not provide evidence that the cumulative incidence of hypothyroidism increased with increasing dose.

Figure IX.G-2. Distribution of Simulation Estimates of Logistic Regression Coefficient: Hypothyroidism



H. Autoimmune (Hashimoto's) Thyroiditis

H.1. Occurrence of Autoimmune (Hashimoto's) Thyroiditis

The primary and alternative definitions for autoimmune (Hashimoto's) thyroiditis were as follows:

- Primary definition: HTDS evaluation or medical records with supporting documentation (625 cases)
- Alternative definition #1: HTDS evaluation or medical records with or without supporting documentation (628 cases)
- Alternative definition #2: Any diagnosis or participant/respondent report (629 cases).

Of the 3440 living evaluable participants, 629 (18.3%) had diagnoses of autoimmune thyroiditis (Table IX.H-1), with all but four based on the HTDS evaluation or medical records with supporting documentation.

Table IX.H-1. Basis for Diagnosis of Autoimmune Thyroiditis, by Sex

	Female		Male		Total	
Basis for Diagnosis	No.	%	No.	%	No.	%
Yes	407	23.3	222	13.1	629	18.3
▪ HTDS evaluation	402	23.0	221	13.1	623	18.1
▪ Medical records with supporting documentation	1	0.1	1	0.1	2	0.1
▪ Medical records *without* supporting documentation	3	0.2	0	--	3	0.1
▪ Participant/respondent report	1	0.1	0	--	1	0.0
No	1333	76.3	1469	86.8	2802	81.5
Unknown	7	0.4	2	0.1	9	0.3
Total	1747	100.0	1693	100.0	3440	100.0

Nine living evaluable participants were classified as "unknown" with regard to diagnosis of autoimmune thyroiditis. These nine did not have medical record or participant/respondent reports of such diagnoses, and did not have a blood draw (8) or had an insufficient quantity of blood drawn to perform the AMA or anti-TPO test (1) and therefore no HTDS evaluation could be made. These nine participants were included as non-cases in analyses of the dose-response for autoimmune thyroiditis.

H.1.a. Additional Outcomes Related to Assay for Anti-Thyroid Immune Response

Late in the course of the study, it was decided to assay anti-thyroid globulin antibody (anti-TG) levels in the serum specimens that had been provided by nearly all study participants and stored in frozen form. The anti-TG test, although not considered to be the best test of autoimmune thyroiditis, provides an additional marker of antithyroid immune response. This made it possible to define additional outcomes of autoimmune thyroiditis based on the combined results of AMA/anti-TPO and anti-TG, or on anti-TG alone. Two additional outcomes were defined:

- Positive antibodies on the AMA/anti-TPO and/or the anti-TG test. A total of 779 living evaluable participants were antibody-positive based on either or both of their AMA/anti-TPO results and their anti-TG results or had a diagnosis of autoimmune thyroiditis based on medical records with supporting documentation. These 779 comprised 22.6% of the 3440 living evaluable participants (Table IX.H-2).

- Positive antibodies on the anti-TG test, without regard to the AMA/anti-TPO results or a diagnosis of autoimmune thryoiditis based on medical records with supporting documentation. A total of 507

living evaluable participants met these criteria and comprised 14.7% of the 3440 living evaluable participants (Table IX.H-3).

Table IX.H-2. Diagnosis of Autoimmune Thyroiditis Based on AMA/anti-TPO and/or anti-TG, or Medical Records with Supporting Documentation, by Sex

	Female		Male		Total	
Autoimmune Thyroiditis	No.	%	No.	%	No.	%
Yes	500	28.6	279	16.5	779	22.6
No	1247	71.4	1414	83.5	2661	77.4
Total	1747	100.0	1693	100.0	3440	100.0

Table IX.H-3. Diagnosis of Autoimmune Thyroiditis Based on anti-TG, or Medical Records with Supporting Documentation, by Sex

	Female		Male		Total	
Autoimmune Thyroiditis	No.	%	No.	%	No.	%
Yes	327	18.7	180	10.6	507	14.7
No	1420	81.3	1513	89.4	2933	85.3
Total	1747	100.0	1693	100.0	3440	100.0

H.1.b. Additional Outcomes Related to Autoimmune Thyroiditis and Hypothyroidism

Two additional outcomes of autoimmune thyroiditis in combination with hypothyroidism were defined to narrow the definition of autoimmune thyroiditis to include both an immune marker for autoimmune thyroid disease and hypothyroidism. These outcomes would represent the most advanced stages of the autoimmune process (hypothyroidism). These additional outcomes were added to determine if a dose-response might be seen with these most advanced stages but missed in the broader category of autoimmune thyroiditis where hypothyroidism had not yet occurred. For this purpose, the diagnoses of autoimmune thyroiditis and hypothyroidism in these additional outcomes were based on the primary definitions, i.e., on the HTDS evaluation or medical records with supporting documentation. The two additional outcomes were:

- Autoimmune thyroiditis (positive AMA/anti-TPO) in participants who also had a diagnosis of hypothyroidism. There were 175 (5.1%) such cases (Table IX.H-4).

- Autoimmune thyroiditis (positive AMA/anti-TPO) in participants who also had a diagnosis of non-iatrogenic, permanent hypothyroidism. This outcome was similar to the first, but excluded those with an iatrogenic cause of hypothyroidism (surgery or ^{131}I therapy) or with transient hypothyroidism. One hundred and sixty-one (4.7%) living evaluable participants met this definition (Table IX.H-5).

Table IX.H-4. Cross-tabulation of Disease Status with Respect to Diagnosis of Autoimmune Thyroiditis in combination with Hypothyroidism, by Sex

Autoimmune Thyroiditis	Female		Male		Total	
with Hypothyroidism	No.	%	No.	%	No.	%
Yes	135	7.7	40	2.4	175	5.1
No	1612	92.3	1653	97.6	3265	94.9
Total	1747	100.0	1693	100.0	3440	100.0

Table IX.H-5. Autoimmune Thyroiditis in Combination with Non-Iatrogenic, Permanent Hypothyroidism, by Sex

Autoimmune Thyroiditis with Non-iatrogenic Hypothyroidism	Female		Male		Total	
	No.	%	No.	%	No.	%
Yes	122	7.0	39	2.3	161	4.7
No	1625	93.0	1654	97.7	3279	95.3
Total	1747	100.0	1693	100.0	3440	100.0

H.2. Analysis of Autoimmune (Hashimoto's) Thyroiditis Risk

H.2.a. Primary Analysis

Of the 625 participants with a diagnosis of autoimmune thyroiditis based on the HTDS examination or medical records with supporting documentation, 43 were out-of-area participants for whom the CIDER program could not calculate dose estimates. The proportions with autoimmune thyroiditis according to the primary and two alternative definitions are shown by sex, dose category and basis for diagnosis in Table IX.H-6.

Table IX.H-6. Diagnoses of Autoimmune (Hashimoto's) Thyroiditis by Sex, Estimated Dose, and Basis for Diagnosis

A. Female

Thyroid Radiation Dose (mGy)	Living Evaluable Female	Primary Definition: Cases Based on HTDS or Med. Rec. with Supporting Documentation		1st Alternative Definition: Cases based on HTDS or Med. Rec. with or without Supporting Documentation		2nd Alternative Definition: Cases Based on Any Diagnosis or Participant or CATI Report	
	No.	No.	%	No.	%	No.	%
Out of Area	125	22	17.6	22	17.6	22	17.6
< 10	182	44	24.2	44	24.2	45	24.7
10-49	320	71	22.2	71	22.2	71	22.2
50-99	313	81	25.9	82	26.2	82	26.2
100-149	220	53	24.1	54	24.5	54	24.5
150-199	126	36	28.6	36	28.6	36	28.6
200-299	139	29	20.9	29	20.9	29	20.9
300-399	144	33	22.9	34	23.6	34	23.6
400-999	171	32	18.7	32	18.7	32	18.7
1000+	7	2	28.6	2	28.6	2	28.6
Total	1747	403	23.1	406	23.2	407	23.3

B. Male

Thyroid Radiation Dose (mGy)	Living Evaluable Male	Primary Definition: Cases Based on HTDS or Med. Rec. with Supporting Documentation		1st Alternative Definition: Cases Based on HTDS or Med. Rec. with or without Supporting Documentation		2nd Alternative Definition: Cases Based on Any Diagnosis or Participant or CATI Report	
	No.	No.	%	No.	%	No.	%
Out of Area	124	21	16.9	21	16.9	21	16.9
< 10	186	26	14.0	26	14.0	26	14.0
10-49	314	40	12.7	40	12.7	40	12.7
50-99	310	47	15.2	47	15.2	47	15.2
100-149	171	17	9.9	17	9.9	17	9.9
150-199	109	12	11.0	12	11.0	12	11.0
200-299	148	18	12.2	18	12.2	18	12.2
300-399	160	20	12.5	20	12.5	20	12.5
400-999	154	20	13.0	20	13.0	20	13.0
1000+	17	1	5.9	1	5.9	1	5.9
Total	1693	222	13.1	222	13.1	222	13.1

Table IX.H-7 displays the numbers of participants with diagnoses of autoimmune thyroiditis when anti-TG was used in addition to AMA/anti-TPO to identify antithyroid immune response, or when anti-TG was used alone.

Table IX.H-7. Additional Disease Outcomes Related to Autoimmune Thyroiditis by Sex and Estimated Dose (cases based on HTDS examination or medical records with supporting documentation only)

A. Female

Thyroid Radiation Dose (mGy)	Living Evaluable Female	HTDS Diagnosis from AMA/anti-TPO and/or anti-TG, or Medical Record with Supporting Documentation		HTDS Diagnosis from anti-TG, or Medical Record with Supporting Documentation	
	No.	No.	%	No.	%
Out of Area	125	25	20.0	13	10.4
< 10	182	62	34.1	45	24.7
10-49	320	91	28.4	57	17.8
50-99	313	89	28.4	49	15.7
100-149	220	72	32.7	57	25.9
150-199	126	42	33.3	27	21.4
200-299	139	35	25.2	24	17.3
300-399	144	38	26.4	28	19.4
400-999	171	44	25.7	25	14.6
1000+	7	2	28.6	2	28.6
Total	1747	500	28.6	327	18.7

B. Male

Thyroid Radiation Dose (mGy)	Living Evaluable Male	HTDS Diagnosis from AMA/anti-TPO and/or anti-TG, or Medical Record with Supporting Documentation		HTDS Diagnosis from anti-TG, or Medical Record with Supporting Documentation	
	No.	No.	%	No.	%
Out of Area	124	26	21.0	20	16.1
< 10	186	33	17.7	26	14.0
10-49	314	51	16.2	29	9.2
50-99	310	58	18.7	35	11.3
100-149	171	21	12.3	13	7.6
150-199	109	17	15.6	10	9.2
200-299	148	24	16.2	18	12.2
300-399	160	24	15.0	16	10.0
400-999	154	24	15.6	13	8.4
1000+	17	1	5.9	0	--
Total	1693	279	16.5	180	10.6

Table IX.H-8 displays the numbers of participants with diagnoses of autoimmune thyroiditis together with diagnoses of hypothyroidism.

Table IX.H-8. Disease Outcomes Related to Autoimmune Thyroiditis with Hypothyroidism by Sex and Estimated Dose (cases based on HTDS examination or medical records with supporting documentation only)

A. Female

Thyroid Radiation Dose (mGy)	Living Evaluable Female	Autoimmune Thyroiditis with Any Hypothyroidism		Autoimmune Thyroiditis with Non-Iatrogenic, Permanent Hypothyroidism	
	No.	No.	%	No.	%
Out of Area	125	10	8.0	9	7.2
< 10	182	14	7.7	12	6.6
10-49	320	25	7.8	22	6.9
50-99	313	27	8.6	23	7.3
100-149	220	11	5.0	10	4.5
150-199	126	10	7.9	10	7.9
200-299	139	14	10.1	13	9.4
300-399	144	13	9.0	12	8.3
400-999	171	10	5.8	10	5.8
1000+	7	1	14.3	1	14.3
Total	1747	135	7.7	122	7.0

B. Male

Thyroid Radiation Dose (mGy)	Living Evaluable Male	Autoimmune Thyroiditis with Any Hypothyroidism		Autoimmune Thyroiditis with Non-Iatrogenic, Permanent Hypothyroidism	
	No.	No.	%	No.	%
Out of Area	124	5	4.0	5	4.0
< 10	186	4	2.2	4	2.2
10-49	314	6	1.9	6	1.9
50-99	310	8	2.6	8	2.6
100-149	171	3	1.8	3	1.8
150-199	109	1	0.9	1	0.9
200-299	148	4	2.7	4	2.7
300-399	160	5	3.1	5	3.1
400-999	154	4	2.6	3	1.9
1000+	17	0	--	0	--
Total	1693	40	2.4	39	2.3

Parameter estimates for the linear dose-response model based on the 3191 in-area participants are shown in Table IX.H-9 below. Based on maximum likelihood analysis of the sex-stratified linear probability model, and using the primary dose estimates, the estimated slope of the dose-response was slightly less than zero (–0.026 per Gy; row 1 of Table IX.H-9) with Bonferroni-adjusted 95% confidence interval ranging from less than –0.057 to 0.044 per Gy, providing no evidence that cumulative incidence increased with increasing dose ($p = 0.82$). The corresponding estimated background rates for diagnosis of benign thyroid nodule were 0.239 with confidence interval (0.212, 0.267) for women and 0.133 with confidence interval (0.109, 0.156) for men. Very similar results were obtained when the model was fit by the method of least squares using ungrouped or grouped data (Table IX.H-9, rows 2 and 3).

Table IX.H-9. Summary of Dose-Response Results for Diagnoses of Autoimmune Thyroiditis

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
1.	Primary definition (HTDS evaluation or medical record with documentation)	Linear	Primary	None	MLE	.239 ± .012 (.212, .267)	.133 ± .010 (.109, .156)	–.026 ± .026 (< –.057, .044)	0.82
2.	Primary definition	Linear	Primary	None	LSU	.240 ± .011 (.214, .266)	.133 ± .011 (.107, .160)	–.029 ± .030 (–.102, .043)	0.83
3.	Primary definition	Linear	Primary	None	LSG	.243 ± .011 (.216, .270)	.137 ± .012 (.109, .164)	–.048 ± .035 (–.131, .035)	0.92
4.	Alternative def. #1 (HTDS or medical record with or without documentation)	Linear	Primary	None	MLE	.241 ± .012 (.213, .269)	.133 ± .010 (.109, .156)	–.025 ± .026 (< –.057, .044)	0.82
5.	Alternative def. #2 (Any diagnosis or participant/respondent report)	Linear	Primary	None	MLE	.242 ± .012 (.214, .269)	.133 ± .010 (.109, .156)	–.026 ± .026 (< –.057, .044)	0.83
6.	Diagnoses based on AMA/anti-TPO and/or anti-TG	Linear	Primary	None	MLE	.300 ± .012 (.270, .329)	.168 ± .011 (.143, .194)	–.039 ± .029 (< –.071, .036)	0.90

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.H-9. Summary of Dose-Response Results for Diagnoses of Autoimmune Thyroiditis (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
7.	Diagnoses based on anti-TG, without regard to AMA/anti-TPO	Linear	Primary	None	MLE	.199 ± .011 (.174, .225)	.107 ± .009 (.087, .128)	–.032 ± .022 (< –.045, .030)	0.90
8.	Autoimmune thyroiditis with any hypothyroidism	Linear	Primary	None	MLE	.077 ± .007 (.060, .094)	.022 ± .005 (.011, .033)	.000 ± .015 (< –.010, >.015)	0.50
9.	Autoimmune thyroiditis with non-iatrogenic, permanent hypothyroidism	Linear	Primary	None	MLE	.070 ± .007 (.053, .086)	.022 ± .004 (.011, .032)	.001 ± .015 (< –.010, .043)	0.48
10.	Primary definition	LQ	Primary	None	LSU	.246 ± .012 (.217, .276)	.140 ± .012 (.110, .170)	Lin: –.090 ± .051 (–.218, .038) Quad: .050 ± .034 (–.035, .134)	Quad: 0.14
11.	Primary definition	Logistic	Primary	None	MLE	.242 (.212, .273)	.132 (.111, .157)	–.22 ± .22 (-.74, .31)	0.84

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.H-9. Summary of Dose-Response Results for Diagnoses of Autoimmune Thyroiditis (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
12.	Primary definition	Linear	Primary	Exclude dose > 1000 mGy	MLE	.241 ± .012 (.212, .269)	.135 ± .011 (.110, .161)	–.038 ± .034 (–.113, .048)	0.86
13.	Primary definition	Linear	Primary	Exclude dose > 400 mGy	MLE	.245 ± .013 (.213, .277)	.134 ± .012 (.105, .162)	–.042 ± .064 (–.191, .116)	0.74
14.	Primary definition	Linear	Primary	Exclude Ok and F/S geostrata	MLE	.238 ± .012 (.209, .267)	.133 ± .010 (.108, .157)	–.025 ± .027 (< –.057, .047)	0.81
15.	Primary definition	Linear	Alt. #1	None	MLE	.238 ± .012 (.210, .267)	.131 ± .010 (.107, .155)	-.016 ± .027 (<-.057, .055)	0.72
16.	Primary definition	Linear	Alt. #2	None	MLE	.240 ± .012 (.212, .268)	.134 ± .010 (.110, .158)	-.030 ± .027 (<-.062, .041)	0.85
17.	Primary definition	Linear	Primary	Include OOA (scoping analysis #1)	MLE	.234 ± .011 (.208, .261)	.135 ± .009 (.113, .157)	-.023 ± .026 (<-.057, .046)	0.80
18.	Primary definition	Linear	Primary	Include OOA (scoping analysis #2)	MLE	.235 ± .011 (.209, .261)	.135 ± .009 (.113, .158)	-.027 ± .026 (<-.057, >.041)	0.84

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

H.2.b. Alternative Definitions for Diagnosis of Autoimmune (Hashimoto's) Thyroiditis

Two alternative definitions for cases of autoimmune thyroiditis were considered. The first alternative added three participants with diagnoses based on medical records without supporting documentation, for a total of 628 (585 in-area, 43 out-of-area) cases. These three cases had estimated doses of 396, 107, and 77 mGy.

The second alternative added a single case based on a report from the participant or his/her CATI respondent, bringing the total to 629 (586 in-area, 43 out-of-area) cases. This case had an estimated dose of 8 mGy.

As shown in rows 4 and 5 of Table IX.H-9, the parameter estimates for the linear dose-response model for these alternative definitions were essentially identical to those obtained in the primary analysis (row 1). In particular there was no evidence for either alternative definition that the cumulative incidence of autoimmune thyroiditis increased with increasing dose.

H.2.b.1. Additional Outcomes Related to Assay for Antithyroid Immune Response

The HTDS diagnoses of autoimmune thyroiditis in the analyses described above were based on the AMA or anti-TPO values that were obtained as part of the participants' HTDS examinations. Since anti-TG measurements were also available, additional analyses were performed to assess the impact of incorporating anti-TG into the diagnostic criterion. Two additional diagnostic criteria were considered. The first required a positive AMA/anti-TPO, a positive anti-TG, or both, or medical records with supporting documentation, and increased the number of cases to 779. As shown in row 6 of Table IX.H-9, the estimated slope of the dose-response for this outcome was less than zero (–0.039 per Gy) with Bonferroni-adjusted 95% confidence interval ranging from less than –0.071 to 0.036 per Gy, providing no evidence that cumulative incidence increased with increasing dose ($p = 0.90$).

The second additional criterion required only a positive anti-TG or medical records with supporting documentation, resulting in a total of 507 cases. As shown in row 7 of Table IX.H-9, the estimated slope of the dose-response for this outcome was less than zero (–0.032 per Gy) with Bonferroni-adjusted 95% confidence interval ranging from less than –0.045 to 0.030 per Gy, providing no evidence that cumulative incidence increased with increasing dose ($p = 0.90$).

H.2.b.2. Additional Outcomes Related to Autoimmune (Hashimoto's) Thyroiditis and Hypothyroidism

Further analyses were made to examine the dose-responses for diagnoses of autoimmune thyroiditis with hypothyroidism. The sex-stratified linear model [1] was fit using the primary criteria for defining cases with both autoimmune thyroiditis and hypothyroidism (HTDS examination or medical record with supporting documentation). Two definitions of the outcome, varying in characteristics of hypothyroidism allowed, were considered (see section IX.H.1.b above). As shown in rows 8 and 9 of Table IX.H-9 above, for both definitions the estimated slope of the sex-stratified linear dose-response model was not significantly greater than zero ($p = 0.50$ and 0.48).

H.2.c. Alternative Dose-Response Functions

As shown in row 10 of Table IX.H-9, the estimated regression coefficient for the dose-squared term in the linear-quadratic dose-response model [5] was 0.050 with Bonferroni-adjusted 95% confidence interval ranging from –0.035 to 0.134. Thus the addition of a quadratic term did not significantly improve the fit of the model ($p = 0.14$).

In the analysis of autoimmune thyroiditis based on the HTDS examination or medical records with supporting documentation, i.e., the primary criterion for defining cases with autoimmune thyroiditis, the regression parameter for the effect of dose in the sex-stratified logistic regression model [2] was estimated as –0.22 with Bonferroni-adjusted 95% confidence interval ranging from –0.74 to 0.31 (Table IX.H-9, row 11). Thus there was no evidence from the logistic regression model that cumulative incidence of autoimmune thyroiditis increased with increasing dose (p = 0.84).

H.2.d. Effect of Excluding Participants in High Dose Categories

When participants with the highest doses were excluded, there was still no evidence that the cumulative incidence of autoimmune thyroiditis increased with increasing dose, as shown in rows 12 and 13 of Table IX.H-9.

H.2.e. Effect of Excluding Okanogan and Ferry/Stevens Geostrata

When Okanogan and Ferry/Stevens geostrata were excluded the estimated slope B was –0.025 per Gy, with Bonferoni-adjusted 95% confidence interval ranging from less than –0.057 to 0.047 per Gy, providing no evidence that the cumulative incidence of autoimmune thyroiditis increased with increasing dose (p = 0.81; row 14 of Table IX.H-9).

H.2.f. Analysis of Autoimmune (Hashimoto's) Thyroiditis in Relation to Alternative Dose Estimates

As shown in rows 15 and 16 of Table IX.H-9, there was no major change in the dose-response results when the alternative dose estimates were used, and in neither case was there evidence that the cumulative incidence increased with increasing dose (p = 0.72 and p = 0.85 for the first and second set of dose estimates, respectively).

H.2.g. Scoping Analysis Regarding Out-of-Area Participants

See section VIII.C.1.a.3 for a description of the scoping analyses that were performed to assess the possible impact of including the 249 out-of-area participants. In neither of the scoping analyses was there evidence that the cumulative incidence increased with increasing dose (p = 0.80 and p = 0.84 for the first and second scoping analyses, respectively; Table IX.H-9, rows 17 and 18).

H.2.h. *Analysis of Autoimmune (Hashimoto's) Thyroiditis in Relation to Alternative Representations of Exposure*

In the analyses by geostratum and by dichotomous exposure variable, the sex and age-adjusted comparisons of cumulative incidence were performed as described in section VIII.C.2.a.2.

H.2.h.1. Analysis by Geostratum

As shown in Table IX.H-10, the proportions of women with autoimmune thyroiditis (by HTDS or medical record with documentation) ranged from 21/75 (28.0%) for the Okanogan geostratum to 35/170 (20.6%) for the Walla Walla County geostratum. For men the proportion ranged from 57/358 (15.9%) to 51/501 (10.2%) for the Benton County and Pasco/Kennewick geostrata, respectively. This heterogeneity

among the nine geostrata was not considered statistically significant (p = 0.073). The percentages with autoimmune thyroiditis were somewhat higher in the Okanogan and Ferry/Stevens geostrata (26.6% for women, 14.2% for men) than in the remaining geostrata (22.8% and 13.0%), but this heterogeneity between combined geostrata was not statistically significant (p = 0.12).

Table IX.H-10. Diagnoses of Autoimmune Thyroiditis Based on the HTDS Evaluation or on Medical Records with Supporting Documentation, by Geostratum and Sex

	Female			Male			Total		
Geostratum	No.	Cases	%	No.	Cases	%	No.	Cases	%
Richland	179	43	24.0	173	24	13.9	352	67	19.0
Pasco/Kennewick	508	106	20.9	501	51	10.2	1009	157	15.6
Benton County	376	92	24.5	358	57	15.9	734	149	20.3
Franklin County	73	19	26.0	76	11	14.5	149	30	20.1
Adams County	165	35	21.2	156	23	14.7	321	58	18.1
Walla Walla (city)	133	35	26.3	131	18	13.7	264	53	20.1
Walla Walla County	170	35	20.6	164	19	11.6	334	54	16.2
Okanogan County	75	21	28.0	64	10	15.6	139	31	22.3
Ferry/Stevens Counties	68	17	25.0	70	9	12.9	138	26	18.8
Total	1747	403	23.1	1693	222	13.1	3440	625	18.2

H.2.h.2. Analysis by Dichotomous Exposure Variable

Of the 1257 participants included in these analyses, 210 (16.7%) had a diagnosis of autoimmune thyroiditis based on the HTDS examination or medical records with supporting documentation (see Table IX.H-11). These included 92/580 (15.9%) in the high exposure group and 118/677 (17.4%) in the low exposure group. The cumulative incidence of autoimmune thyroiditis was not significantly higher in the high exposure group (p = 0.86).

Table IX.H-11. Diagnoses of Autoimmune Thyroiditis Based on HTDS or Medical Record with Supporting Documentation, by Exposure Group and Sex

	Female			Male			Total		
Exposure Group	No.	Cases	%	No.	Cases	%	No.	Cases	%
Low	351	70	19.9	326	48	14.7	677	118	17.4
High	298	60	20.1	282	32	11.3	580	92	15.9
Total	649	130	20.0	608	80	13.2	1257	210	16.7

H.2.i. Confounding and Effect Modification

As described in section VIII above, additional sex-stratified logistic regression models were investigated to examine the possibility that the primary dose-response results might be influenced by confounding, and to search for factors that might modify a radiation dose-response. These analyses were based on the primary definition of Hashimoto's thyroiditis, i.e., those based on an HTDS diagnosis or on medical records with documented diagnoses, and on the primary dose estimates. Table IX.H-12 displays results for models including sex, age at first exposure to Hanford ^{131}I (prenatal, or < 180 days), age at HTDS examination, estimated dose from the NTS, history of any cancer other than thyroid, and HTDS interview type.

Note that sex was not analyzed as a possible confounder since its effect was already adjusted for in the sex-stratified model. None of the other factors in Table IX.H-12 appears to be a confounder: for none does the adjusted estimate of the regression coefficient differ markedly from the unadjusted estimate.

Therefore, it does not appear that omitting these factors introduces any important bias in the dose-response results.

The analyses of effect modification address the question of whether the dose-response might vary according to other characteristics of the study participants. This was tested by comparing the estimated regression coefficients for the groups defined by each covariate. As shown in Table IX.H-12, the regression coefficients did not differ significantly between the groups defined by any of the covariates, suggesting that none of them was a significant modifier of a radiation dose-response for Hashimoto's thyroiditis.

Table IX.H-12. Confounding and Effect Modification by Sex, Age at Exposure or HTDS Examination, Estimated Thyroid Dose from NTS, History of Any Cancer Other than Thyroid, and Interview Type: Autoimmune Thyroiditis

		Estimated Dose-Response Coefficient (per Gy)				
Covariate (0=No, 1=Yes)	Yes / Total	Unadjusted	Adjusted for Confounding	Including Effect Modification Group 0	Including Effect Modification Group 1	P
Female?	1622 / 3191	–.215 ± .220 (–.742, .313)	Not Applicable	–.223 ± .354 (–1.11, .660)	–.210 ± .282 (–.913, .494)	.98
Prenatal exposure?	1034 / 3191	–.215 ± .220 (–.742, .313)	–.255 ± .224 (–.832, .322)	–.183 ± .252 (–.849, .483)	–.491 ± .473 (–1.74, .756)	.56
1st exposure before age 180 days?	1478 / 3191	–.215 ± .220 (–.742, .313)	–.211 ± .222 (–.782, .360)	–.781 ± .395 (–1.82, .260)	.059 ± .251 (–.602, .720)	.071
Age at exam > 50?	2001 / 3191	–.215 ± .220 (–.742, .313)	–.310 ± .229 (–.900, .280)	–.659 ± .484 (–1.94, .617)	–.193 ± .261 (–.882, .496)	.38
NTS ^{131}I dose > 5.3 mGy?	1567 / 3189	–.214 ± .220 (–.741, .313)	–.222 ± .226 (–.803, .360)	–.354 ± .306 (–1.16, .454)	–.056 ± .328 (–.922, .810)	.51
History of any cancer other than thyroid?	248 / 3186	–.206 ± .220 (–.733, .320)	–.201 ± .221 (–.769, .367)	–.145 ± .234 (–.762, .472)	–.661 ± .762 (–2.67, 1.35)	.49
Expanded In-Person Interview?	1212 / 3191	–.215 ± .220 (–.742, .313)	–.273 ± .226 (–.855, .309)	–.783 ± .382 (–1.79, .226)	.015 ± .264 (–.681, .712)	.084

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

Tables IX.H-13 and IX.H-14 display similar results from analyses including history of medical or dental x-ray exposure or of occupational exposures as potential confounding or effect modifying factors. The estimates of the regression coefficient calculated with adjustment for confounding are all close to the unadjusted estimates. Moreover the adjusted estimates all remained less than zero. Thus there was no evidence that a confounding effect of any of these covariates has obscured a positive dose-response for autoimmune thyroiditis.

There is no evidence of any statistically significant effect modification by any of the covariates in Tables IX.H-13 and IX.H-14.

Table IX.H-13. Confounding and Effect Modification by History of Medical and Dental Radiation Exposures: Autoimmune Thyroiditis

Have You Ever Had: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted	Adjusted for Confounding	Including Effect Modification Group 0	Including Effect Modification Group 1	P
CAT scan of the upper body?	775 / 3149	−.209 ± .221 (−.738, .320)	−.211 ± .221 (−.780, .358)	−.225 ± .246 (−.874, .423)	−.149 ± .509 (−1.49, 1.19)	.89
Diagnostic x-rays of the head?	1191 / 3155	−.207 ± .222 (−.739, .325)	−.213 ± .223 (−.787, .360)	−.197 ± .274 (−.920, .526	−.245 ± .381 (−1.25, .761)	.92
Diagnostic x-rays of the neck?	966 / 3167	−.204 ± .221 (−.732, .325)	−.199 ± .222 (−.770, .373)	−.369 ± .301 (−1.16, .424)	.014 ± .318 (−.824, .852)	.38
Diagnostic x-rays of chest or upper body, including mammograms?	2821 / 3173	−.199 ± .220 (−.725, .327)	−.198 ± .220 (−.765, .368)	.048 ± .773 (−1.99, 2.09)	−.219 ± .230 (−.827, .388)	.74
Diagnostic x-rays of the stomach or mid-back?	692 / 3120	−.224 ± .224 (−.760, .311)	−.229 ± .224 (−.806, .348)	−.238 ± .250 (−.898, .422)	−.192 ± .503 (−1.52, 1.14)	.94
Barium enema?	825 / 3159	−.235 ± .223 (−.768, .298)	−.236 ± .223 (−.810, .337)	−.065 ± .248 (−.719, .588)	−.811 ± .498 (−2.13, .504)	.17
Upper GI?	1146 / 3177	−.233 ± .222 (−.764, .298)	−.227 ± .222 (−.798, .344)	−.156 ± .271 (−.872, .559)	−.360 ± .382 (−1.37, .648)	.66
Intravenous pyelogram?	398 / 3157	−.223 ± .222 (−.756, .309)	−.213 ± .222 (−.784, .359)	−.201 ± .234 (−.818, .415)	−.311 ± .700 (−2.16, 1.53)	.88
Fluoroscopy of the upper body?	246 / 3161	−.224 ± .221 (−.755, .306)	−.221 ± .221 (−.791, .349)	−.191 ± .228 (−.792, .409)	−.639 ± .886 (−2.98, 1.70)	.62
Nuclear scan (excluding thyroid scan)?	217 / 3162	−.185 ± .219 (−.709, .339)	−.178 ± .219 (−.742, .386)	−.155 ± .224 (−.746, .436)	−.579 ± .968 (−3.13, 1.97)	.66
History of any cancer other than thyroid?	248 / 3186	−.206 ± .220 (−.733, .320)	−.201 ± .221 (−.769, .367)	−.145 ± .234 (−.762, .472)	−.661 ± .762 (−2.67, 1.35)	.49
Dental x-rays that did not usually include a lead shield over the neck area?	1648 / 3191	−.215 ± .220 (−.742, .313)	−.216 ± .220 (−.783, .351)	.092 ± .294 (−.684, .869)	−.554 ± .336 (−1.44, .332)	.15

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

Table IX.H-14. Confounding and Effect Modification by Occupational History: Autoimmune Thyroiditis

Have You Ever Worked in Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted	Adjusted for Confounding	Including Effect Modification Group 0	Including Effect Modification Group 1	P
Any metal industry?	238 / 3191	−.215 ± .220 (−.742, .313)	−.217 ± .220 (−.785, .351)	−.144 ± .222 (−.730, .442)	−1.75 ± 1.15 (−4.78, 1.27)	.14
Any nuclear facility?	371 / 3191	−.215 ± .220 (−.742, .313)	−.226 ± .222 (−.798, .346)	−.196 ± .242 (−.835, .442)	−.377 ± .561 (−1.86, 1.10)	.77
Any other industry or occupation where you may have been exposed to radioactive materials or x-rays?	442 / 3191	−.215 ± .220 (−.742, .313)	−.210 ± .220 (−.777, .358)	−.079 ± .233 (−.693, .534)	−1.26 ± .800 (−3.37, .848)	.12
Any of the above industries or occupations?	892 / 3191	−.215 ± .220 (−.742, .313)	−.223 ± .221 (−.792, .346)	−.015 ± .255 (−.688, .658)	−.746 ± .448 (−1.93, .435)	.14

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

Table IX.H-15 displays the results of analyses of possible confounding or effect modification by smoking variables. There was no evidence that the dose-response was significantly confounded by either smoking variable, or that there was a dose-response that differed significantly according to smoking history.

Table IX.H-15. Confounding and Effect Modification by Smoking: Autoimmune Thyroiditis

Have You Ever Smoked Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted	Adjusted for Confounding	Including Effect Modification Group 0	Including Effect Modification Group 1	P
Cigarettes (unfiltered or filtered)?	1854 / 3183	−.210 ± .220 (−.736, .317)	−.214 ± .220 (−.781, .352)	−.664 ± .393 (−1.70, .373)	.009 ± .257 (−.668, .686)	.15
Any of cigarettes, cigar or pipe?	1900 / 3183	−.210 ± .220 (−.736, .317)	−.214 ± .220 (−.781, .352)	−.609 ± .394 (−1.65, .430)	−.021 ± .258 (−.701, .658)	.21

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

H.2.j. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for autoimmune thyroiditis are shown in Figure IX.H-1 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in that figure are calculated at the 98.33% confidence level, i.e., are adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates.

The point estimate of the slope was greater than zero for only 13 of the 100 realizations, and the confidence interval included zero for all 100 realizations. Also shown in Figure IX.H-1 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean and mean of each participant's 100 dose realizations. In summary, for none of the 100 realizations of the estimated doses was there a statistically significant dose-response, and for most realizations the estimated slope was less than 0.

Figure IX.H-1. Plot of Estimated Slope and 95% CI by Dose Realization: Autoimmune Thyroiditis

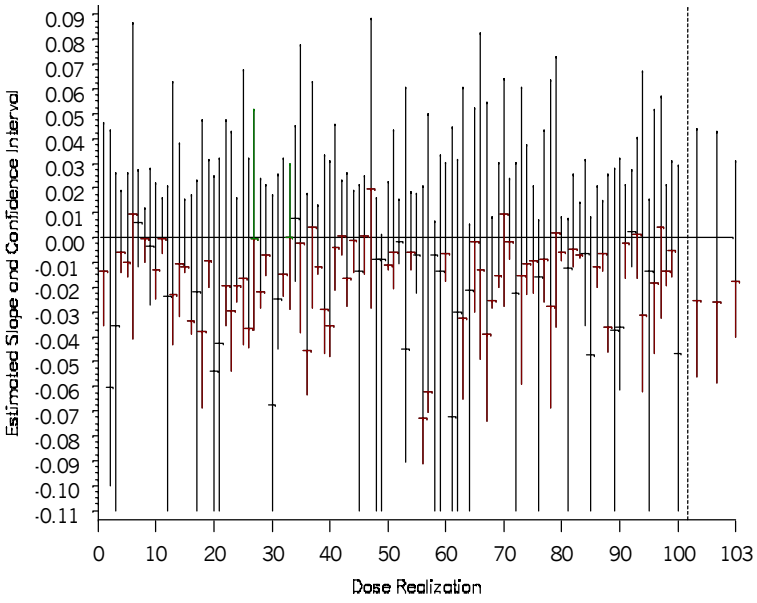
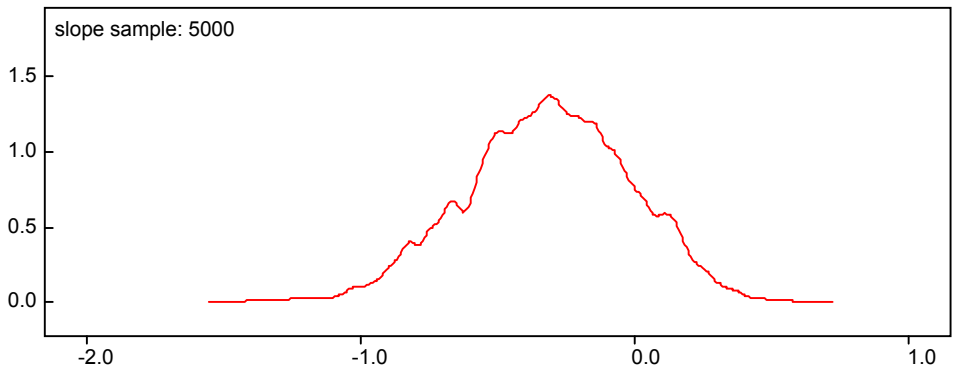


Figure IX.H-2 displays the distribution of the 5000 estimates of the logistic regression coefficient obtained by the simulation procedures described in section VIII.C.2.b.3 above. It is evident from the figure that most of the estimates were between about –1.5 and 0.5. The estimate was less than or equal to 0 for 4453 of the 5000 replications, implying an empirical one-tailed p-value of 0.89. The median estimate was –0.36, and the upper and lower percentiles corresponding to the Bonferroni-adjusted 95% confidence interval ranging from –1.35 to 0.32. These may be compared to the estimates of –0.22 (with confidence interval ranging from –0.74 to 0.31) obtained using the median dose estimates without adjustment for uncertainty. Thus, this method of adjusting the estimated logistic regression coefficient for the uncertainty in the dose estimates did not provide evidence that the cumulative incidence of autoimmune thyroiditis increased with increasing dose.

Figure IX.H-2. Distribution of Simulation Estimates of Logistic Regression Coefficient: Autoimmune Thyroiditis



I. Graves Disease

I.1. Occurrence of Graves Disease

The primary and alternative definitions of Graves disease were as follows:

- Primary definition: HTDS evaluation or medical records with supporting documentation (34 cases)
- Alternative definition #1: HTDS evaluation or medical records with or without supporting documentation (37 cases)
- Alternative definition #2: Any diagnosis or participant/respondent report (50 cases).

Thirty-four (1.0%) living evaluable participants had a diagnosis of Graves Disease based on the HTDS evaluation or on medical records with supporting documentation (Table IX.I-1). Three (0.1%) living evaluable participants had a diagnosis of Graves Disease based on medical records without supporting documentation, and an additional thirteen (0.4%) were based on a participant or his/her CATI respondent report.

Table IX.I-1. Basis for Diagnosis of Graves Disease, by Sex

	Female		Male		Total	
Basis for Diagnosis	No.	%	No.	%	No.	%
Yes	37	2.1	13	0.8	50	1.5
▪ HTDS evaluation	5	0.3	2	0.1	7	0.2
▪ Medical records with supporting documentation	23	1.3	4	0.2	27	0.8
▪ Medical records *without* supporting documentation	3	0.2	0	--	3	0.1
▪ Participant/respondent report	6	0.3	7	0.4	13	0.4
No	1698	97.2	1673	98.8	3371	98.0
Unknown	12	0.7	7	0.4	19	0.6
Total	1747	100.0	1693	100.0	3440	100.0

Nineteen living evaluable participants were classified as "unknown" with regard to diagnosis of Graves disease. These 19 did not have medical records or participant/respondent reports of such diagnoses, and did not have an HTDS evaluation due to lack of blood draw (8) or diagnosis of hyperthyroidism with unknown etiology (potentially Graves) (11). These 19 participants were included as non-cases in analyses of the dose-response for Graves disease.

I.2. Analysis of Graves Disease Risk

I.2.a. Primary Analysis

Of the 34 living evaluable participants with a diagnosis of Graves disease based on the HTDS examination or medical records with supporting documentation, two were out-of-area participants for whom the CIDER program could not calculate dose estimates. The proportions with Graves disease are shown by sex, dose category and basis for diagnosis in Table IX.I-2.

Table IX.I-2. Diagnoses of Graves Disease by Sex, Estimated Dose, and Basis for Diagnosis

A. Female

Thyroid Radiation Dose (mGy)	Living Evaluable Female	Primary Definition: Cases Based on HTDS or Med. Rec. with Supporting Documentation		1st Alternative Definition: Cases Based on HTDS or Med. Rec. with or without Supporting Documentation		2nd Alternative Definition: Cases Based on Any Diagnosis or Participant or CATI Report	
	No.	No.	%	No.	%	No.	%
Out of Area	125	2	1.6	2	1.6	2	1.6
< 10	182	5	2.7	7	3.8	7	3.8
10-49	320	4	1.3	4	1.3	5	1.6
50-99	313	3	1.0	4	1.3	7	2.2
100-149	220	2	0.9	2	0.9	2	0.9
150-199	126	0	--	0	--	0	--
200-299	139	4	2.9	4	2.9	5	3.6
300-399	144	2	1.4	2	1.4	3	2.1
400-999	171	6	3.5	6	3.5	6	3.5
1000+	7	0	--	0	--	0	--
Total	1747	28	1.6	31	1.8	37	2.1

B. Male

Thyroid Radiation Dose (mGy)	Living Evaluable Female	Primary Definition: Cases Based on HTDS or Med. Rec. with Supporting Documentation		1st Alternative Definition: Cases Based on HTDS or Med. Rec. with or without Supporting Documentation		2nd Alternative Definition: Cases Based on Any Diagnosis or Participant or CATI Report	
	No.	No.	%	No.	%	No.	%
Out of Area	124	0	--	0	--	0	--
< 10	186	1	0.5	1	0.5	1	0.5
10-49	314	1	0.3	1	0.3	2	0.6
50-99	310	2	0.6	2	0.6	2	0.6
100-149	171	0	--	0	--	1	0.6
150-199	109	1	0.9	1	0.9	2	1.8
200-299	148	0	--	0	--	2	1.4
300-399	160	1	0.6	1	0.6	3	1.9
400-999	154	0	--	0	--	0	--
1000+	17	0	--	0	--	0	--
Total	1693	6	0.4	6	0.4	13	0.8

Parameter estimates for the linear dose-response model based on the 3191 in-area participants are shown in Table IX.I-3 below. Based on the maximum likelihood analysis of the sex-stratified linear probability model, the estimated slope B was slightly less than zero (–0.001 per Gy) with Bonferroni-adjusted 95% CI ranging from less than –0.002 to 0.024 per Gy, providing no evidence that cumulative incidence increased with increasing dose (one-tailed p=0.56; row 1 of Table IX.I-3). The corresponding estimated background rates for diagnosis of Graves disease were 0.016 with confidence interval (0.008, 0.025) for women and 0.004 with confidence interval (0, 0.009) for men.

As shown in rows 2 and 3 of Table IX.I-3, generally similar results were obtained when the model was fit by the method of least squares. The estimates of the slope were slightly but not significantly greater than zero (p = 0.26 and 0.13 for ungrouped and grouped data, respectively).

Table IX.I-3. Summary of Dose-Response Results for Diagnoses of Graves Disease

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
1.	Primary definition (HTDS evaluation or medical record with documentation)	Linear	Primary	None	MLE	.016 ± .004 (.008, .025)	.004 ± .002 (0*, .009)	–.001 ± .009 (< –.002, .024)	0.56
2.	Primary definition	Linear	Primary	None	LSU	.015 ± .003 (.008, .022)	.003 ± .003 (0*, .010)	.005 ± .008 (–.014, .024)	0.26
3.	Primary definition	Linear	Primary	None	LSG	.014 ± .003 (.007, .021)	.002 ± .003 (0*, .009)	.010 ± .009 (–.012, .032)	0.13
4.	Alternative def. #1 (HTDS or medical record with or without documentation)	Linear	Primary	None	MLE	.018 ± .004 (.009, .027)	.004 ± .002 (0*, .009)	–.002 ± .009 (NE, .020)	0.64
5.	Alternative def. #2 (Any diagnosis or participant/respondent report)	Linear	Primary	None	MLE	.021 ± .004 (.011, .032)	.008 ± .003 (.001, .016)	.001 ± .013 (< –.004, .034)	0.48

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model. "0*" indicates that the lower confidence limit for a background rate was less than 0.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.I-3. Summary of Dose-Response Results for Diagnoses of Graves Disease (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
6.	Primary definition	LQ	Primary	None	LSU	.014 ± .003 (.007, .022)	.002 ± .003 (0*, .010)	Lin: .014 ± .013 (–.020, .047) Quad: –.007 ± .009 (–.029, .015)	Quad: 0.43
7.	Primary definition	Logistic	Primary	None	MLE	.015 (.008, .026)	.004 (.001, .010)	.42 ± .65 (-1.13, 1.96)	0.28
8.	Primary definition	Linear	Primary	Exclude dose > 1000 mGy	MLE	.016 ± .004 (.007, .025)	.004 ± .002 (0*, .008)	.001 ± .009 (<-.005, .029)	0.44
9.	Primary definition	Linear	Primary	Exclude dose > 400 mGy	MLE	.014 ± .003 (.006, .022)	.004 ± .002 (0*, .010)	.0005 ± .012 (–.023, .037)	0.48
10.	Primary definition	Linear	Primary	Exclude Ok and F/S geostrata	MLE	.015 ± .004 (.006, .024)	.002 ± .002 (0*, .006)	.003 ± .008 (< –.001, .027)	0.36
11.	Primary definition	Linear	Alt. #1	None	MLE	.016 ± .003 (.008, .025)	.004 ± .002 (0*, .009)	–.002 ± .008 (NE, .015)	0.70

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model. "0*" indicates that the lower confidence limit for a background rate was less than 0.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.I-3. Summary of Dose-Response Results for Diagnoses of Graves Disease

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
12.	Primary definition	Linear	Alt. #2	None	MLE	.015 ± .004 (.007, .024)	.003 ± .002 (0*, .008)	.003 ± .008 (< –.002, .026)	0.34
13.	Primary definition	Linear	Primary	Include OOA (scoping analysis #1)	MLE	.016 ± .003 (.008, .024)	.004 ± .002 (0*, .008)	.000 ± .008 (< –.002, >.025)	0.50
14.	Primary definition	Linear	Primary	Include OOA (scoping analysis #2)	MLE	.016 ± .003 (.008, .024)	.004 ± .002 (0*, .008)	–.0003 ± .008 (< –.002, >.024)	0.51

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model. "0*" indicates that the lower confidence limit for a background rate was less than 0.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

I.2.b. Alternative Definitions for Diagnosis of Graves Disease

Two alternative definitions for cases of Graves disease were considered. The first alternative added three cases with diagnoses based on medical records without supporting documentation, for a total of 37 cases. The second alternative criterion for defining cases of Graves disease added another 13 cases based solely on a report from the participant or his/her CATI respondent, for a total of 50 cases. As shown in rows 4 and 5 of Table IX.I-3 above, for neither of these alternative definitions was there any evidence that the cumulative incidence of Graves disease increased significantly with increasing dose ($p = 0.64$ and $p = 0.48$ for the first and second alternative criteria respectively).

I.2.c. Alternative Dose-Response Functions

As shown in row 6 of Table IX.I-3, the estimated regression coefficient for the dose-squared term in the linear-quadratic dose-response model [5] was –0.007 with Bonferroni-adjusted 95% confidence interval ranging from –0.029 to 0.015. Thus the addition of a quadratic term did not significantly improve the fit of the model ($p = 0.43$).

In the analysis of Graves disease based on the HTDS examination or medical records with supporting documentation, i.e., the primary definition of Graves disease, the regression parameter for the effect of dose in the sex-stratified logistic regression model was estimated as 0.42 with Bonferroni-adjusted 95% confidence interval ranging from –1.13 to 1.96 (row 7 of Table IX.I-3). Thus the cumulative incidence of Graves disease did not increase significantly with increasing dose ($p = 0.28$).

I.2.d. Effect of Excluding Participants in High Dose Categories

As shown in row 8 of Table IX.I-3, when participants with estimated dose > 1000 mGy were excluded, the estimated slope B was not significantly greater than zero (0.001 per Gy, with Bonferroni-adjusted 95% confidence interval ranging from less than –0.005 to 0.029 per Gy; $p = 0.44$). Similar results were obtained when participants with estimated dose > 400 mGy were excluded (Table IX.I-3, row 9).

I.2.e. Effect of Excluding Okanogan and Ferry/Stevens Geostrata

As shown in row 10 of Table IX.I-3, when Okanogan and Ferry/Stevens geostrata were excluded, the estimated slope B was not significantly greater than zero (0.003 per Gy, with Bonferroni-adjusted 95% confidence interval ranging from less than –0.001 to 0.027 per Gy; $p = 0.36$).

I.2.f. Analysis of Graves Disease in Relation to Alternative Dose Estimates

For neither set of alternative dose estimates did the cumulative incidence increase significantly with increasing dose ($p = 0.70$ and $p = 0.34$ for the first and second set of alternative dose estimates, respectively; Table IX.I-3, rows 11 and 12).

I.2.g. Scoping Analysis Regarding Out-of-Area Participants

See section VIII.C.1.a.3 for a description of the scoping analyses that were performed to assess the possible impact of including the 249 out-of-area participants. As summarized in rows 13 and 14 of Table IX.I-3, in neither scoping analysis was there any evidence that the cumulative incidence of Graves Disease

increased with increasing dose (p = 0.50 and p = 0.51 for the first and second scoping analyses, respectively).

I.2.h. Analysis of Graves Disease in Relation to Alternative Representations of Exposure

In the analyses by geostratum and by dichotomous exposure variable, the sex and age-adjusted comparisons of cumulative incidence were performed as described in section VIII.C.2.a.2.

I.2.h.1. Analysis by Geostratum

There were too few participants (34) with diagnoses of Graves disease (from the HTDS examination or medical records with documentation) for a definitive conclusion regarding heterogeneity among the geostrata (see Table IX.I-4). The absence of significant heterogeneity (p = 0.43) was not strong evidence against the possibility that the cumulative incidence of Graves disease might vary among geostrata. The percentages with Graves disease were somewhat higher in the Okanogan and Ferry/Stevens geostrata (2.8% for women, 1.5% for men) than in the remaining geostrata (1.5% and 0.3%), but this heterogeneity between combined geostrata was also not statistically significant (p = 0.13).

Table IX.I-4. Diagnoses of Graves Disease Based on the HTDS Evaluation or on Medical Records with Supporting Documentation, by Geostratum and Sex

	Female			Male			Total		
Geostratum	No.	Cases	%	No.	Cases	%	No.	Cases	%
Richland	179	3	1.7	173	0	--	352	3	0.9
Pasco/Kennewick	508	10	2.0	501	1	0.2	1009	11	1.1
Benton County	376	5	1.3	358	1	0.3	734	6	0.8
Franklin County	73	1	1.4	76	1	1.3	149	2	1.3
Adams County	165	3	1.8	156	1	0.6	321	4	1.2
Walla Walla (city)	133	1	0.8	131	0	--	264	1	0.4
Walla Walla County	170	1	0.6	164	0	--	334	1	0.3
Okanogan County	75	2	2.7	64	1	1.6	139	3	2.2
Ferry/Stevens Counties	68	2	2.9	70	1	1.4	138	3	2.2
Total	1747	28	1.6	1693	6	0.4	3440	34	1.0

I.2.h.2. Analysis by Dichotomous Exposure Variable

Only 13 (1.0%) of the 1257 participants included in these analyses had a diagnosis of Graves disease based on the HTDS examination or medical records with supporting documentation (see Table IX.I-5). These included 7/580 (1.2%) in the high exposure group and 6/677 (0.9%) in the low exposure group. The cumulative incidence of Graves disease was not significantly greater in the high exposure group (p = 0.24).

Table IX.I-5. Diagnoses of Graves Disease based on HTDS or Medical Record with Supporting Documentation, by Exposure Group and Sex

	Female			Male			Total		
Exposure Group	No.	Cases	%	No.	Cases	%	No.	Cases	%
Low	351	5	1.4	326	1	0.3	677	6	0.9
High	298	6	2.0	282	1	0.4	580	7	1.2
Total	649	11	1.7	608	2	0.3	1257	13	1.0

I.2.i. Confounding and Effect Modification

As described in section VIII above, additional sex-stratified logistic regression models were investigated to examine the possibility that the primary dose-response results might be influenced by confounding, and to search for factors that might modify a radiation dose-response. These analyses were based on the primary definition of Graves Disease; those based on an HTDS diagnosis or on medical records with documented diagnoses, and on the primary dose estimates. Table IX.I-6 displays results for models including sex, age at first exposure to Hanford ^{131}I (prenatal, or < 180 days), age at HTDS examination, estimated dose from the NTS, history of any cancer other than thyroid, and HTDS interview type.

Note that sex was not analyzed as a possible confounder since its effect was already adjusted for in the sex-stratified model. None of the other factors in Table IX.I-6 appears to be a confounder: for none does the adjusted estimate of the regression coefficient differ markedly from the unadjusted estimate. Therefore, it does not appear that omitting these factors introduces any important bias in the dose-response results.

The analyses of effect modification address the question of whether the dose-response might vary according to other characteristics of the study participants. This was tested by comparing the estimated regression coefficients for the groups defined by each covariate. As shown in Table IX.I-6, the regression coefficients did not differ significantly between the groups defined by any of the covariates, suggesting that none of them was a significant modifier of a radiation dose-response for Graves disease.

Table IX.I-6. Confounding and Effect Modification by Sex, Age at Exposure or HTDS Examination, Estimated Thyroid Dose from NTS, History of Any Cancer Other than Thyroid, and Interview Type: Graves Disease

		Estimated Dose-Response Coefficient (per Gy)				
Covariate (0=No, 1=Yes)	Yes / Total	Unadjusted	Adjusted for Confounding	Including Effect Modification Group 0	Including Effect Modification Group 1	P
Female?	1622 / 3191	.415 ± .646 (–1.13, 1.96)	Not Applicable	–1.78 ± 2.83 (–8.85, 5.29)	.635 ± .622 (–.918, 2.19)	.32
Prenatal exposure?	1034 / 3191	.415 ± .646 (–1.13, 1.96)	.370 ± .659 (–1.33, 2.07)	.440 ± .704 (–1.42, 2.30)	.038 ± 1.67 (–4.37, 4.44)	.82
1st exposure before age 180 days?	1478 / 3191	.415 ± .646 (–1.13, 1.96)	.393 ± .643 (–1.26, 2.05)	2.03 ± 1.01 (–.633, 4.69)	-.771 ± 1.32 (–4.26, 2.72)	.071
Age at exam > 50?	2001 / 3191	.415 ± .646 (–1.13, 1.96)	.493 ± .625 (–1.12, 2.10)	.267 ± 1.11 (–2.66, 3.20)	.628 ± .765 (–1.39, 2.64)	.78
NTS ^{131}I dose > 5.3 mGy?	1567 / 3189	.415 ± .646 (–1.13, 1.96)	.319 ± .691 (–1.46, 2.10)	.770 ± .679 (–1.02, 2.56)	–2.08 ± 2.52 (–8.72, 4.56)	.19
History of any cancer other than thyroid?	248 / 3186	.415 ± .646 (–1.13, 1.96)	.508 ± .682 (–1.25, 2.26)	.508 ± .682 (–1.29, 2.31)	0.0 ± 1115 (–2943, 2943)	1.0
Expanded In-Person Interview?	1212 / 3191	.415 ± .646 (–1.13, 1.96)	.499 ± .651 (–1.18, 2.18)	1.40 ± .992 (–1.22, 4.01)	-.275 ± 1.27 (–3.64, 3.09)	.26

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

Tables IX.I-7 and IX.I-8 display similar results from analyses including history of medical or dental x-ray exposure or of occupational exposures as potential confounding or effect modifying factors. The estimates of the regression coefficient calculated with adjustment for confounding are all close to the unadjusted estimates. Thus there was no evidence that omitting these factors introduces any important bias in the dose-response results for Graves Disease.

There is no evidence of any statistically significant effect modification by any of the covariates in Tables IX.I-7 and IX.I-8, with the possible exception of history of diagnostic x-rays of the chest or upper back, including mammograms (Table IX.I-7). However the regression parameter for the 352 participants without such histories is extremely negative, –275, with an extremely wide confidence interval (–817, 267), since only two participants in this group had diagnoses of Graves disease (both women with doses less than 10 mGy). Therefore the p-value of 0.002 for effect modification must be interpreted cautiously. It is noteworthy that the regression parameter for the larger group of participants with histories of chest or upper body diagnostic x-rays or mammograms (0.534 with confidence interval ranging from –1.07 to 2.14) differs little from the overall estimate of 0.414 with confidence interval (–1.13, 1.96).

Table IX.I-7. Confounding and Effect Modification by History of Medical and Dental Radiation Exposures: Graves Disease

Have You Ever Had: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy) Unadjusted	Adjusted for Confounding	Including Effect Modification Group 0	Including Effect Modification Group 1	P
CAT scan of the upper body?	775 / 3149	.423 ± .650 (–1.13, 1.98)	.433 ± .654 (–1.25, 2.12)	.421 ± .716 (–1.47, 2.31)	.500 ± 1.66 (–3.88, 4.88)	.97
Diagnostic x-rays of the head?	1191 / 3155	.423 ± .642 (–1.11, 1.96)	.421 ± .635 (–1.22, 2.06)	.246 ± .852 (–2.00, 2.50)	.721 ± .973 (–1.84, 3.29)	.72
Diagnostic x-rays of the neck?	966 / 3167	.376 ± .670 (–1.23, 1.98)	.329 ± .669 (–1.39, 2.05)	.212 ± 1.14 (–2.79, 3.21)	.392 ± .804 (–1.73, 2.51)	.90
Diagnostic x-rays of chest or upper body, including mammograms?	2821 / 3173	.414 ± .646 (–1.13, 1.96)	.439 ± .641 (–1.21, 2.09)	–275 ± 205 (–817, 267)	.534 ± .608 (–1.07, 2.14)	.002
Diagnostic x-rays of the stomach or mid-back?	692 / 3120	.427 ± .642 (–1.11, 1.96)	.412 ± .644 (–1.25, 2.07)	.362 ± .714 (–1.52, 2.24)	.669 ± 1.48 (–3.23, 4.57)	.86
Barium enema?	825 / 3159	.368 ± .673 (–1.24, 1.98)	.375 ± .675 (–1.36, 2.11)	.685 ± .699 (–1.16, 2.53)	–.650 ± 1.69 (–5.10, 3.80)	.42
Upper GI?	1146 / 3177	.417 ± .645 (–1.13, 1.96)	.409 ± .647 (–1.26, 2.08)	.536 ± .744 (–1.43, 2.50)	.120 ± 1.25 (–3.17, 3.41)	.77
Intravenous pyelogram?	398 / 3157	.376 ± .670 (–1.23, 1.98)	.382 ± .664 (–1.33, 2.09)	.386 ± .681 (–1.41, 2.18)	.318 ± 2.87 (–7.26, 7.90)	.98
Fluoroscopy of the upper body?	246 / 3161	.381 ± .667 (–1.22, 1.98)	.393 ± .667 (–1.33, 2.11)	.336 ± .713 (–1.54, 2.22)	1.13 ± 2.35 (–5.08, 7.34)	.76
Nuclear scan (excluding thyroid scan)?	217 / 3162	.418 ± .643 (–1.12, 1.96)	.417 ± .640 (–1.23, 2.07)	.382 ± .659 (–1.36, 2.12)	2.11 ± 4.02 (–8.48, 12.7)	.69
History of any cancer other than thyroid cancer?	248 / 3186	.415 ± .646 (–1.13, 1.96)	.508 ± .682 (–1.25, 2.26)	.508 ± .682 (–1.29, 2.31)	0.0 ± 1115 (–2943, 2943)	1.0
Dental x-rays that did not usually include a lead shield over the neck area?	1648 / 3191	.415 ± .646 (–1.13, 1.96)	.403 ± .654 (–1.28, 2.09)	.271 ± .914 (–2.14, 2.68)	.558 ± .91 (–1.83, 2.95)	.82

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

Table IX.I-8. Confounding and Effect Modification by Occupational History: Graves Disease

Have You Ever Worked in Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted	Adjusted for Confounding	Including Effect Modification Group 0	Including Effect Modification Group 1	P
Any metal industry?	238 / 3191	.415 ± .646 (–1.13, 1.96)	.423 ± .643 (–1.23, 2.08)	.329 ± .687 (–1.48, 2.14)	3.33 ± 3.19 (–5.09, 11.8)	.40
Any nuclear facility?	371 / 3191	.415 ± .646 (–1.13, 1.96)	.390 ± .674 (–1.35, 2.13)	.399 ± .699 (–1.44, 2.24)	.292 ± 2.40 (–6.04, 6.63)	.97
Any other industry or occupation where you may have been exposed to radioactive materials or x-rays?	442 / 3191	.415 ± .646 (–1.13, 1.96)	.434 ± .631 (–1.19, 2.06)	.242 ± .804 (–1.88, 2.36)	.801 ± .93 (–1.64, 3.24)	.66
Any of the above industries or occupations?	892 / 3191	.415 ± .646 (–1.13, 1.96)	.402 ± .649 (–1.27, 2.07)	.361 ± .811 (–1.78, 2.50)	.480 ± 1.07 (–2.35, 3.31)	.93

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

Table IX.I-9 displays the results of analyses of possible confounding or effect modification by smoking variables. There was no evidence that the dose-response was significantly confounded by either smoking variable, or that there was a dose-response that differed significantly according to smoking history.

Table IX.I-9. Confounding and Effect Modification by Smoking: Graves Disease

Have You Ever Smoked Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficent (per Gy)				
		Unadjusted	Adjusted for Confounding	Including Effect Modification Group 0	Including Effect Modification Group 1	P
Cigarettes (unfiltered or filtered)?	1854 / 3183	.407 ± .658 (–1.17, 1.98)	.439 ± .666 (–1.28, 2.15)	.665 ± 1.05 (–2.11, 3.44)	.300 ± .90 (–2.08, 2.68)	.79
Any of cigarettes, cigar or pipe?	1900 / 3183	.407 ± .658 (–1.17, 1.98)	.441 ± .666 (–1.28, 2.16)	.663 ± 1.052 (–2.11, 3.44)	.303 ± .90 (–2.08, 2.69)	.80

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

I.2.j. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for Graves disease are shown in Figure IX.I-1 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in that figure are calculated at the 98.33% confidence level, i.e., are adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates. The point estimate of the slope was greater than 0 for 48 of the 100 realizations, and the confidence interval included 0 for all 100 realizations. Also shown in Figure IX.I-1 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean and mean of each participant's 100 dose realizations. In summary, for none of the 100 realizations of the estimated doses was there a statistically significant dose-response, and for about half of the realizations the estimated slope was less than 0.

Figure IX.I-1. Plot of estimated Slope and 95% CI by Dose Realization: Graves Disease

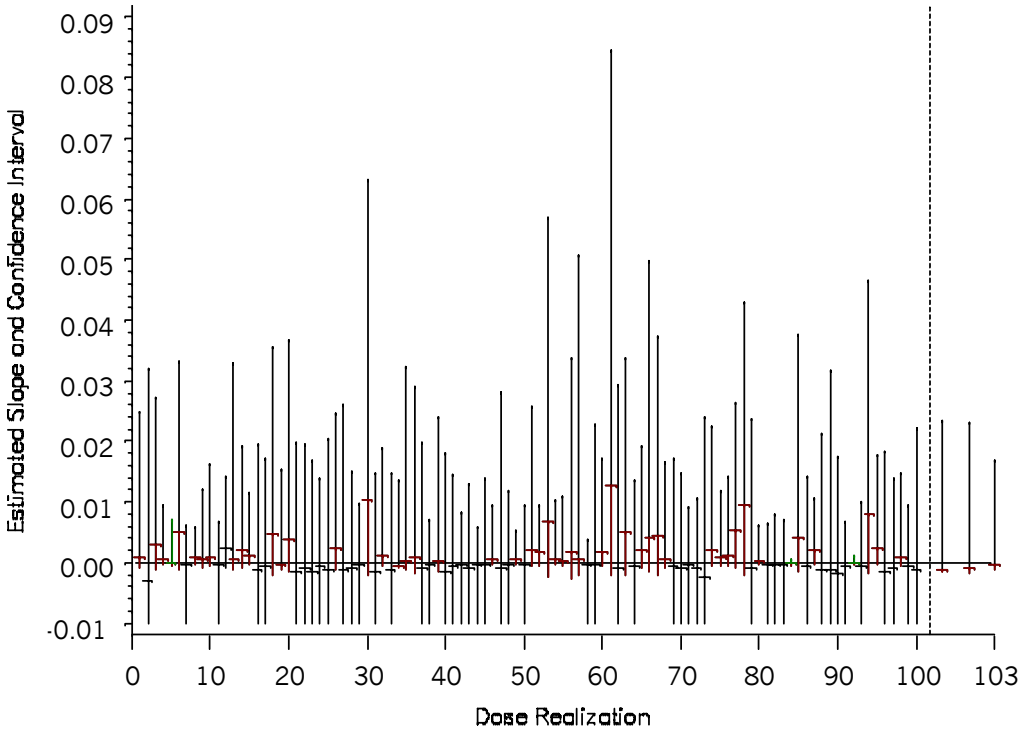
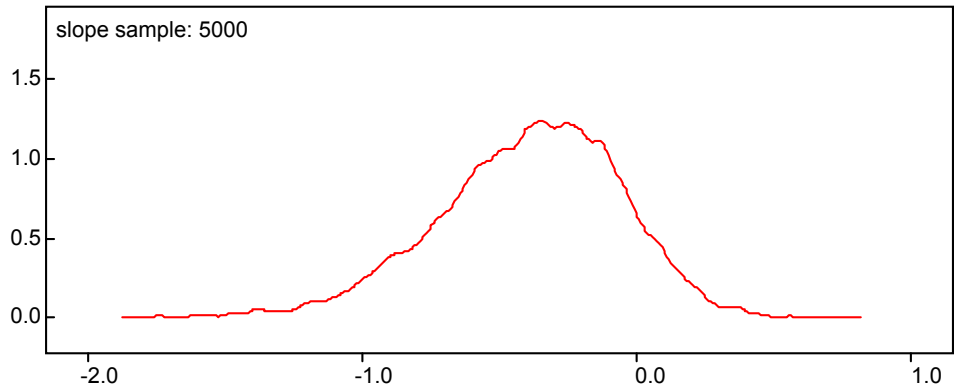


Figure IX.I-2 displays the distribution of the 5000 estimates of the logistic regression coefficient obtained by the simulation procedures described in section VIII.C.2.b.3 above. It is evident from the figure that most of the estimates were between about –1.5 and 0.5. The estimate was less than or equal to 0 for 2068 of the 5000 replications, implying an empirical one-tailed p-value of 0.41. The median estimate was 0.21, and the upper and lower percentiles corresponding to the Bonferroni-adjusted 95% confidence interval were –3.21 and 1.87. These may be compared to the estimate of 0.42 with confidence interval (–1.13, 1.96) obtained using the median dose estimate without adjustment for uncertainty. Thus, this method of adjusting the estimated logistic regression coefficient for the uncertainty in the dose estimates did not provide evidence that the cumulative incidence of Graves disease increased with increasing dose.

Figure IX.I-2. Distribution of Simulation Estimates of Logistic Regression Coefficient: Graves Disease



J. Autoimmune Thyroid Disease

J.1. Occurrence of Autoimmune Thyroid Disease

Autoimmune thyroid disease was defined by diagnosis of autoimmune (Hashimoto's) thyroiditis or Graves disease. The primary and alternative definitions of autoimmune thyroid disease were as follows:

- Primary definition: Diagnosis of autoimmune thyroiditis or Graves disease based on HTDS evaluation or medical records with supporting documentation (659 cases)
- Alternative definition #1: HTDS evaluation or medical records with or without supporting documentation (663 cases)
- Alternative definition #2: Any diagnosis or participant/respondent report (674 cases).

Of the 3440 living evaluable participants, 659 (19.2%) had a diagnosis of autoimmune thyroid disease based on the HTDS evaluation or medical records with supporting documentation (Table IX.J-1). These included 625 with autoimmune (Hashimoto's) thyroiditis (see section IX.H) and 34 others with diagnoses of Graves disease (see section IX.I). An additional 4 (0.1%) living evaluable participants had a diagnosis of autoimmune thyroid disease based on medical records without supporting documentation (three with autoimmune thyroiditis, one with Graves disease). Eleven other participants (0.3%) were based on a report by the participant or his/her CATI respondent (one with autoimmune thyroiditis, 10 with Graves disease).

Table IX.J-1. Basis for Diagnosis of Autoimmune Thyroid Disease, by Sex

	Female		Male		Total	
Basis for Diagnosis	No.	%	No.	%	No.	%
Yes	442	25.3	232	13.7	674	19.6
▪ HTDS Evaluation	421	24.1	226	13.3	647	18.8
▪ Medical Records with supporting documentation	10	0.6	2	0.1	12	0.3
▪ Medical Records *without* supporting documentation	4	0.2	0	--	4	0.1
▪ Participant/respondent report	7	0.4	4	0.2	11	0.3
No	1296	74.2	1454	85.9	2750	79.9
Unknown	9	0.5	7	0.4	16	0.5
Total	1747	100.0	1693	100.0	3440	100.0

Sixteen living evaluable participants were classified as "unknown" with regard to diagnosis of autoimmune thyroid disease. These sixteen did not have medical record or participant/respondent reports of such diagnoses, and did not have an HTDS evaluation due to lack of a blood draw (N=8), insufficient amount of blood drawn to obtain the antibody level (N=1), and a diagnosis of hyperthyroidism with an uncertain etiology (potentially Graves) (N=7). These sixteen participants were included as non-cases in analyses of the dose-response for autoimmune thyroid disease.

J.2. Analysis of Autoimmune Thyroid Disease Risk

J.2.a. Primary Analysis

Of the 659 living evaluable participants with a diagnosis of autoimmune thyroid disease based on the HTDS examination or medical records with supporting documentation, 45 were out-of-area participants. The proportions with autoimmune thyroid disease are shown by sex, dose category and basis for diagnosis in Table IX.J-2.

Table IX.J-2. Diagnoses of Autoimmune Thyroid Disease by Sex, Estimated Dose, and Basis for Diagnosis

A. Female

Thyroid Radiation Dose (mGy)	Living Evaluable Female	Primary Definition Cases Based on HTDS or Med. Rec. with Supporting Documentation		1st Alternative Definition: Cases Based on HTDS or Med. Rec. with or without Supporting Documentation		2nd Alternative Definition: Cases Based on Any Diagnosis or Participant or CATI Report	
	No.	No.	%	No.	%	No.	%
Out of Area	125	24	19.2	24	19.2	24	19.2
< 10	182	49	26.9	49	26.9	50	27.5
10-49	320	75	23.4	75	23.4	76	23.8
50-99	313	84	26.8	86	27.5	89	28.4
100-149	220	55	25.0	56	25.5	56	25.5
150-199	126	36	28.6	36	28.6	36	28.6
200-299	139	33	23.7	33	23.7	34	24.5
300-399	144	35	24.3	36	25.0	37	25.7
400-999	171	38	22.2	38	22.2	38	22.2
1000+	7	2	28.6	2	28.6	2	28.6
Total	1747	431	24.7	435	24.9	442	25.3

B. Male

Thyroid Radiation Dose (mGy)	Living Evaluable Male	Primary Definition: Cases Based on HTDS or Med. Rec. with Supporting Documentation		1st Alternative Definition: Cases Based on HTDS or Med. Rec. with or without Supporting Documentation		2nd Alternative Definition Cases Based on Any Diagnosis or Participant or CATI Report	
	No.	No.	%	No.	%	No.	%
Out of Area	124	21	16.9	21	16.9	21	16.9
< 10	186	27	14.5	27	14.5	27	14.5
10-49	314	41	13.1	41	13.1	41	13.1
50-99	310	49	15.8	49	15.8	49	15.8
100-149	171	17	9.9	17	9.9	18	10.5
150-199	109	13	11.9	13	11.9	14	12.8
200-299	148	18	12.2	18	12.2	20	13.5
300-399	160	21	13.1	21	13.1	21	13.1
400-999	154	20	13.0	20	13.0	20	13.0
1000+	17	1	5.9	1	5.9	1	5.9
Total	1693	228	13.5	228	13.5	232	13.7

Since nearly all of the cases of autoimmune thyroid disease were in fact autoimmune thyroiditis, it was to be expected that dose-response results for these two disease outcomes would be quite similar. Parameter estimates for the linear dose-response model based on the 3191 in-area participants are shown in Table IX.J-3 below. Based on maximum likelihood analysis of the sex-stratified linear probability model, and using the primary dose estimates, the estimated slope B was slightly less than zero (–0.024 per Gy; row 1 of Table IX.J-3) with Bonferroni-adjusted 95% confidence interval ranging from less than –0.058 to 0.048, providing no evidence that cumulative incidence of autoimmune thyroid disease increased with increasing dose ($p = 0.80$). The corresponding estimated background rates for diagnosis of autoimmune thyroid disease were 0.255 with confidence interval (0.227, 0.283) for women and 0.136 with confidence interval (0.112, 0.160) for men. Very similar results were obtained when the model was fit by the method of least squares using ungrouped or grouped data (Table IX.J-3, rows 2 and 3).

Table IX.J-3. Summary of Dose-Response Results for Diagnoses of Autoimmune Thyroid Disease

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
1.	Primary definition (HTDS evaluation or medical record with documentation)	Linear	Primary	None	MLE	.255 ± .012 (.227, .283)	.136 ± .010 (.112, .160)	–.024 ± .027 (<-.058, .048)	0.80
2.	Primary definition	Linear	Primary	None	LSU	.255 ± .011 (.229, .282)	.136 ± .011 (.109, .163)	–.024 ± .031 (–.098, .049)	0.79
3.	Primary definition	Linear	Primary	None	LSG	.257 ± .011 (.230, .285)	.139 ± .012 (.111, .167)	–.038 ± .036 (–.123, .047)	0.86
4.	Alternative def. #1 (HTDS or medical record with or without documentation)	Linear	Primary	None	MLE	.258 ± .012 (.229, .286)	.136 ± .010 (.112, .160)	–.024 ± .027 (<-.058, .048)	0.80
5.	Alternative def. #2 (Any diagnosis or participant/respondent report)	Linear	Primary	None	MLE	.262 ± .012 (.234, .291)	.139 ± .010 (.115, .163)	–.026 ± .028 (-<.059, .047)	0.81

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is less than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.J-3 Summary of Dose-Response Results for Diagnoses of Autoimmune Thyroid Disease (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
6.	Primary definition	LQ	Primary	None	LSU	.261 ± .012 (.231, .290)	.142 ± .012 (.112, .172)	Lin: –.076 ± .052 (–.207, .054) Quad: .043 ± .035 (–.044, .129)	Quad: 0.22
7.	Primary definition	Logistic	Primary	None	MLE	.256 (.227, .288)	.135 (.114, .160)	–.17 ± .21 (-.68, .34)	0.79
8.	Primary definition	Linear	Primary	Exclude dose > 1000 mGy	MLE	.256 ± .012 (.227, .285)	.138 ±.011 (.113, .163)	-.031 ± .035 (<-.109, .057)	0.81
9.	Primary definition	Linear	Primary	Exclude dose > 400 mGy	MLE	.259 ± .014 (.226, .292)	.138 ± .012 (.109, .166)	-.041 ± .065 (-.191, .119)	0.74
10.	Primary definition	Linear	Primary	Exclude Ok and F/S geostrata	MLE	.252 ± .012 (.223, .282)	.134 ± .010 (.110, .159)	-.020 ± .028 (<-.058, .054)	0.76
11.	Primary definition	Linear	Alt. #1	None	MLE	.255 ± .012 (.226, .284)	.136 ± .010 (.112, .160)	-.021 ± .028 (<-.058, .051)	0.76

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is less than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.J-3. Summary of Dose-Response Results for Diagnoses of Autoimmune Thyroid Disease (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
12.	Primary definition	Linear	Alt. #2	None	MLE	.255 ± .012 (.226, .284)	.136 ± .010 (.112, .161)	-.023 ± .029 (<-.064, .050)	0.78
13.	Primary definition	Linear	Primary	Include OOA (scoping analysis #1)	MLE	.250 ± .011 (.223, .277)	.138 ± .009 (.115, .161)	-.021 ± .027 (<-.059, >.051)	0.77
14.	Primary definition	Linear	Primary	Include OOA (scoping analysis #2)	MLE	.251 ± .011 (.224, .278)	.139 ± .009 (.116, .161)	-.024 ± .027 (<-.059, >.046)	0.81

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is less than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

J.2.b. Alternative Definitions for Diagnosis of Autoimmune Thyroid Disease

Two alternative definitions for cases of autoimmune thyroid disease were considered. The first alternative added four cases with diagnoses based on medical records without supporting documentation, for a total of 663 cases. The second added another 11 cases based solely on a report from the participant or his/her CATI respondent, for a total of 674 cases. As shown in rows 4 and 5 of Table IX.J-3 above, the parameter estimates for the linear dose-response model were essentially identical to those obtained in the primary analysis. In particular there was no evidence in either the primary or alternative analyses that the cumulative incidence of autoimmune thyroid disease increased with increasing dose.

J.2.c. Alternative Dose-Response Functions

As shown in row 6 of Table IX.J-3, the estimated regression coefficient for the dose-squared term in the linear-quadratic dose-response model [5] was 0.043 with Bonferroni-adjusted 95% confidence interval ranging from –0.044 to 0.129. Thus the addition of a quadratic term did not significantly improve the fit of the model ($p = 0.22$).

In the analysis of autoimmune thyroid disease based on the HTDS examination or medical records with supporting documentation, i.e., the primary criterion for defining cases with autoimmune thyroid disease, the regression parameter for the effect of dose in the sex-stratified logistic regression model [2] was estimated as –0.17 with Bonferroni-adjusted 95% confidence interval ranging from –0.68 to 0.34 (Table IX.J-3, row 7). Thus there was no evidence from the logistic regression model that cumulative incidence of autoimmune thyroid disease increased with increasing dose ($p = 0.79$).

J.2.d. Effect of Excluding Participants in High Dose Categories

The results were essentially unchanged if participants in the high dose categories were excluded. As shown in row 8 of Table IX.J-3, if participants with estimated doses over 1000 mGy were excluded, the estimated slope of the sex-stratified linear dose-response model was less than zero (–0.031 per Gy) with Bonferroni-adjusted 95% confidence interval ranging from less than –0.109 to 0.057 per Gy. Thus there was no evidence that the cumulative incidence of autoimmune thyroid disease increased with increasing dose ($p = 0.81$). Similar results were obtained if participants with estimated doses exceeding 400 mGy were excluded ($p = 0.74$; Table IX.J-3, row 9).

J.2.e. Effect of Excluding Okanogan and Ferry/Stevens Geostrata

As shown in row 10 of Table IX.J-3, if the Okanogan and Ferry/Stevens geostrata were excluded, the estimated slope increased slightly from –0.024 to –0.020 per Gy, but there was no evidence that the cumulative incidence of autoimmune thyroid disease increased with increasing dose ($p = 0.76$).

J.2.f. Analysis of Autoimmune Thyroid Disease in Relation to Alternative Dose Estimates

As shown in rows 11 and 12 of Table IX.J-3, for neither set of alternative dose estimates was there any evidence that the cumulative incidence of autoimmune thyroid disease increased with increasing dose ($p = 0.76$ and $p = 0.78$ for the first and second dose set estimates, respectively).

J.2.g. Scoping Analysis Regarding Out-of-Area Participants

See section VIII.C.1.a.3 for a description of the scoping analyses that were performed to assess the possible impact of including the 249 out-of-area participants. As shown in rows 13 and 14 of Table IX.J-3 above, for neither of the scoping analyses was there any evidence that the cumulative incidence of autoimmune thyroid disease increased with increasing dose ($p = 0.77$ and $p = 0.81$ for the first and second scoping analyses, respectively).

J.2.h. Analysis of Autoimmune Thyroid Disease in Relation to Alternative Representations of Exposure

In the analyses by geostratum and by dichotomous exposure variable, the sex and age-adjusted comparisons of cumulative incidence were performed as described in section VIII.C.2.a.2.

J.2.h.1. Analysis by Geostratum

As shown in Table IX.J-4, among the entire 3440 living evaluable participants, the proportions with autoimmune thyroid disease ranged from 23/75 (30.7% in the Okanogan County geostratum) to 36/170 (21.2%, Walla Walla County) for women, and from 11/64 (17.2%, Okanogan County) to 52/501 (10.4%, Pasco/Kennewick) for men ($p = 0.083$ for heterogeneity among the nine geostrata). In particular the percentages with autoimmune thyroid disease were somewhat higher in the Okanogan and Ferry/Stevens geostrata (29.4% for women, 15.7% for men) than in the remaining geostrata (24.3% and 13.3%, respectively; $p = 0.048$). Since it was likely that participants in the Okanogan and Ferry/Stevens geostrata tended to have lower thyroid doses from Hanford's ^{131}I than those in other geostrata, it does not appear that these differences can be attributed to an effect of Hanford's ^{131}I.

Table IX.J-4. Diagnoses of Autoimmune Thyroid Disease based on the HTDS Evaluation or on Medical Records with Supporting Documentation, by Geostratum and Sex

	Female			Male			Total		
Geostratum	No.	Cases	%	No.	Cases	%	No.	Cases	%
Richland	179	46	25.7	173	24	13.9	352	70	19.9
Pasco/Kennewick	508	116	22.8	501	52	10.4	1009	168	16.7
Benton County	376	97	25.8	358	58	16.2	734	155	21.1
Franklin County	73	20	27.4	76	12	15.8	149	32	21.5
Adams County	165	38	23.0	156	24	15.4	321	62	19.3
Walla Walla (city)	133	36	27.1	131	18	13.7	264	54	20.5
Walla Walla County	170	36	21.2	164	19	11.6	334	55	16.5
Okanogan County	75	23	30.7	64	11	17.2	139	34	24.5
Ferry/Stevens Counties	68	19	27.9	70	10	14.3	138	29	21.0
Total	1747	431	24.7	1693	228	13.5	3440	659	19.2

J.2.h.2. Analysis by Dichotomous Exposure Variable

A total of 223 (17.7%) of the 1257 participants included in these analyses had a diagnosis of autoimmune thyroid disease based on the HTDS examination or medical records with supporting documentation (see Table IX.J-5). These included 99/580 (17.1%) in the high exposure group and 124/677 (18.3%) in the low exposure group. The cumulative incidence of autoimmune thyroid disease was not significantly higher in the high exposure group ($p = 0.80$).

Table IX.J-5. Diagnoses of Autoimmune Thyroid Disease based on the HTDS evaluation or on Medical Records with Supporting Documentation, by Exposure Group and Sex

	Female			Male			Total		
Exposure Group	No.	Cases	%	No.	Cases	%	No.	Cases	%
Low	351	75	21.4	326	49	15.0	677	124	18.3
High	298	66	22.1	282	33	11.7	580	99	17.1
Total	649	141	21.7	608	82	13.5	1257	223	17.7

J.2.i. Confounding and Effect Modification

As described in section VIII above, additional sex-stratified logistic regression models were investigated to examine the possibility that the primary dose-response results might be influenced by confounding, and to search for factors that might modify a radiation dose-response. These analyses were based on the primary definition of autoimmune thyroid disease, i.e., those based on an HTDS diagnosis or on medical records with documented diagnoses, and on the primary dose estimates. Table IX.J-6 displays results for models including sex, age at first exposure to Hanford ^{131}I (prenatal, or < 180 days), age at HTDS examination, estimated dose from the NTS, history of any cancer other than thyroid disease, and HTDS interview type.

Note that sex was not analyzed as a possible confounder since its effect was already adjusted for in the sex-stratified model. None of the other factors in Table IX.J-6 appears to be a confounder: for none does the adjusted estimate of the regression coefficient differ markedly from the unadjusted estimate. Therefore, it does not appear that omitting these factors introduces any important bias in the dose-response results.

The analyses of effect modification address the question of whether the dose-response might vary according to other characteristics of the study participants. This was tested by comparing the estimated regression coefficients for the groups defined by each covariate. As shown in Table IX.J-6, the regression coefficients did not differ significantly between the groups defined by any of the covariates, suggesting that none of them was a significant modifier of a radiation dose-response for autoimmune thyroid disease.

Table IX.J-6. Confounding and Effect Modification by Sex, Age at Exposure or HTDS Examination, Estimated Thyroid Dose from NTS, History of Any Cancer Other than Thyroid, and Interview Type: Autoimmune Thyroid Disease

		Estimated Dose-Response Coefficient (per Gy)				
Covariate (0=No, 1=Yes)	Yes / Total	Unadjusted	Adjusted for Confounding	Including Effect Modification Group 0	Including Effect Modification Group 1	P
Female?	1622 / 3191	–.170 ± .213 (–.680, .341)	Not Applicable	–.259 ± .353 (–1.14, .622)	–.116 ± .268 (–.785, .553)	.74
Prenatal exposure?	1034 / 3191	–.170 ± .213 (–.680, .341)	–.211 ± .217 (–.770, .347)	–.134 ± .244 (–.777, .509)	–.467 ± .461 (–1.68, .748)	.52
1st exposure before age 180 days?	1478 / 3191	–.170 ± .213 (–.680, .341)	–.167 ± .214 (–.719, .385)	–.550 ± .377 (–1.55, .445)	.019 ± .249 (–.638, .676)	.21
Age at exam > 50?	2001 / 3191	–.170 ± .213 (–.680, .341)	–.250 ± .221 (–.818, .319)	–.574 ± .457 (–1.78, .633)	–.136 ± .253 (–.805, .532)	.39
NTS ^{131}I dose > 5.3 mGy?	1567 / 3189	–.169 ± .213 (–.679, .342)	–.185 ± .219 (–.749, .379)	–.239 ± .291 (–1.01, .529)	–.114 ± .330 (–.984, .756)	.78
History of any cancer other than thyroid?	248 / 3186	–.161 ± .213 (–.671, .348)	–.154 ± .214 (–.704, .397)	–.094 ± .227 (–.692, .504)	–.661 ± .762 (–2.67, 1.35)	
Expanded In-Person Interview?	1212 / 3191	–.170 ± .213 (–.680, .341)	–.218 ± .218 (–.781, .345)	–.595 ± .364 (–1.55, .365)	.002 ± .261 (–.688, .691)	.18

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

Tables IX.J-7 and IX.J-8 display similar results from analyses including history of medical or dental x-ray exposure or of occupational exposures as potential confounding or effect modifying factors. The estimates of the regression coefficient calculated with adjustment for confounding are all close to the unadjusted estimates. Moreover the adjusted estimates all remained less than zero. Thus there was no evidence that a confounding effect of any of these covariates has obscured a positive dose-response for autoimmune disease.

There is no evidence of any statistically significant effect modification by any of the covariates in Tables IX.J-7 and IX.J-8.

Table IX.J-7. Confounding and Effect Modification by History of Medical and Dental Radiation Exposures: Autoimmune Thyroid Disease

Have You Ever Had: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy) Unadjusted	Adjusted for Confounding	Including Effect Modification Group 0	Including Effect Modification Group 1	P
CAT scan of the upper body?	775 / 3149	−.164 ± .214 (−.676, .348)	−.166 ± .214 (−.717, .385)	−.181 ± .238 (−.808, .447)	−.100± .495 (−1.41, 1.21)	.88
Diagnostic x-rays of the head?	1191 / 3155	−.160 ± .215 (−.674, .354)	−.166 ± .215 (−.720, .388)	−.168 ± .266 (−.870, .534)	−.162 ± .366 (−1.13, .803)	.99
Diagnostic x-rays of the neck?	966 / 3167	−.164 ± .214 (−.676, .349)	−.162 ± .215 (−.715, .392)	−.343 ± .294 (−1.12, .432)	.057 ± .306 (−.749, .864)	.35
Diagnostic x-rays of chest or upper body, including mammograms?	2821 / 3173	−.154 ± .213 (−.663, .355)	−.152 ± .213 (−.700, .397)	−.091 ± .787 (−2.17, 1.99)	−.156 ± .221 (−.740, .427)	.94
Diagnostic x-rays of the stomach or mid-back?	692 / 3120	−.176 ± .216 (−.694, .342)	−.182 ± .216 (−.739, .376)	−.192 ± .241 (−.829, .444)	−.136 ± .489 (−1.43, 1.15)	.92
Barium enema?	825 / 3159	−.194 ± .216 (−.711, .323)	−.195 ± .216 (−.751, .361)	−.007 ± .240 (−.640, .626)	−.821 ± .485 (−2.10, .458)	.12
Upper GI?	1146 / 3177	−.186 ± .215 (−.700, .327)	−.181 ± .214 (−.733, .371)	−.101 ± .262 (−.791, .590)	−.333 ± .371 (−1.31, .646)	.61
Intravenous Pyelogram?	398 / 3157	−.182 ± .215 (−.698, .333)	−.171 ± .215 (−.725, .383)	−.159 ± .226 (−.755, .438)	−.285 ± .687 (−2.10, 1.53)	.86
Fluoroscopy of the upper body?	246 / 3161	−.183 ± .215 (−.697, .331)	−.179 ± .215 (−.731, .374)	−.157 ± .221 (−.740, .427)	−.487 ± .846 (−2.72, 1.74)	.70
Nuclear scan (excluding thyroid scan)?	217 / 3162	−.141 ± .212 (−.648, .367)	−.134 ± .212 (−.680, .412)	−.115 ± .217 (−.687, .457)	−.482 ± .947 (−2.98, 2.02)	.70
Dental x-rays that did not usually include a lead shield over the neck area?	1648 / 3191	−.170 ± .213 (−.680, .341)	−.172 ± .213 (−.721, .377)	.111 ± .286 (−.643, .866)	−.484 ± .325 (−1.34, .372)	.17

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

Table IX.J-8. Confounding and Effect Modification by Occupational History: Autoimmune Thyroid Disease

Have You Ever Worked in Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted	Adjusted for Confounding	Including Effect Modification		P
				Group 0	Group 1	
Any metal industry?	238 / 3191	–.170 ± .213 (–.680, .341)	–.171 ± .213 (–.720, .379)	–.111 ± .216 (–.680, .458)	–1.42 ± 1.08 (–4.27, 1.43)	.21
Any nuclear facility?	371 / 3191	–.170 ± .213 (–.680, .341)	–.184 ± .215 (–.739, .371)	–.152 ± .234 (–.769, .465)	–.351 ± .550 (–1.80, 1.10)	.74
Any other industry or occupation where you may have been exposed to radioactive materials or x-rays?	442 / 3191	–.170 ± .213 (–.680, .341)	–.162 ± .213 (–.711, .387)	–.059 ± .227 (–.658, .540)	–.895 ± .710 (–2.77, .977)	.22
Any of the above industries or occupations?	892 / 3191	–.170 ± .213 (–.680, .341)	–.178 ± .214 (–.729, .373)	.014 ± .248 (–.641, .669)	–.654 ± .429 (–1.79, .477)	.16

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

Table IX.J-9 displays the results of analyses of possible confounding or effect modification by smoking variables. There was no evidence that the dose-response was significantly confounded by either smoking variable, or that there was a dose-response that differed significantly according to smoking history.

Table IX.J-9. Confounding and Effect Modification by Smoking: Autoimmune Thyroid Disease

Have You Ever Smoked Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted	Adjusted for Confounding	Including Effect Modification		P
				Group 0	Group 1	
Cigarettes (unfiltered or filtered)?	1854 / 3183	–.167 ± .213 (–.677, .344)	–.170 ± .213 (–.719, .379)	–.563 ± .378 (–1.56, .433)	.028 ± .251 (–.634, .691)	.19
Any of cigarettes, cigar or pipe?	1900 / 3183	–.167 ± .213 (–.677, .344)	–.170 ± .213 (–.719, .379)	–.510 ± .378 (–1.51, .488)	–.001 ± .252 (–.665, .663)	.26

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

J.2.j. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for autoimmune thyroid disease are shown in Figure IX.J-1 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in that figure are calculated at the 98.33% confidence level, i.e., are adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates. While the point estimate of the slope is greater than 0 for only 16 of the 100 realizations, the confidence interval includes 0 for all 100 realizations. Also shown in Figure IX.J-1 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean and mean of each participant's 100 dose realizations. In summary, for none of the 100 realizations of the estimated doses was there a statistically significant dose-response, and for most of the realizations the estimated slope was less than 0.

Figure IX.J-1. Plot of Estimated Slope and 95% CI by Dose Realization: Autoimmune Thyroid Disease

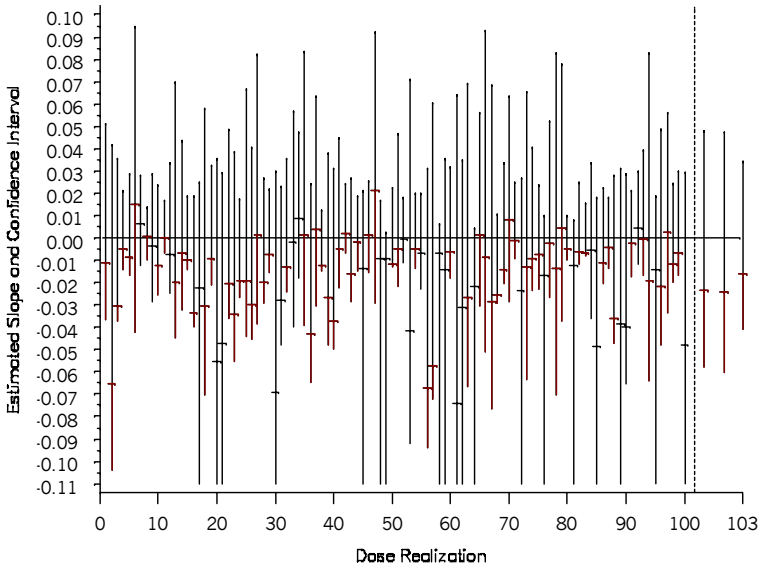
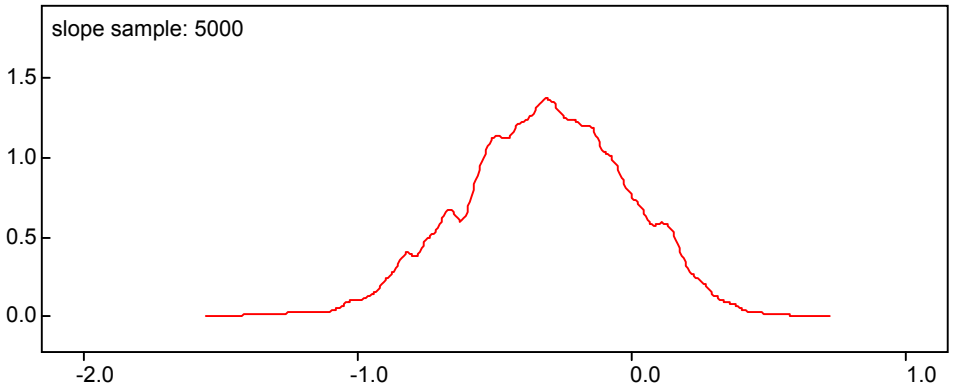


Figure IX.J-2 displays the distribution of the 5000 logistic regression coefficient estimates obtained by the simulation procedures described in section VIII.C.2.b.3 above. It is evident from the figure that most of the estimates were between about –1.5 and 0.5. The estimate was less than or equal to 0 for 4226 of the 5000 replications, implying an empirical one-tailed p-value of 0.85. The median estimate was –0.31, and the upper and lower percentiles corresponding to the Bonferroni-adjusted 95% confidence interval were –1.08 and 0.35. These may be compared to the estimate of –0.17 with confidence interval (–0.68, 0.34) obtained using the median dose estimates without adjustment for uncertainty. Thus, this method of adjusting the estimated logistic regression coefficient for the uncertainty in the dose estimates did not provide evidence that the cumulative incidence of autoimmune thyroid disease increased with increasing dose.

Figure IX.J-2. Distribution of Simulation Estimates of Logistic Regression Coefficient: Autoimmune Thyroid Disease



K. Hyperthyroidism

K.1. Occurrence of Hyperthyroidism

The primary and alternative definitions for hyperthyroidism were as follows:

- Primary definition: HTDS evaluation or medical records with supporting documentation (161 cases)
- Alternative definition #1: HTDS evaluation or medical records with or without supporting documentation (175 cases)
- Alternative definition #2: Any diagnosis or participant/respondent report (196 cases).

There were 161 (4.7%) cases of hyperthyroidism based on the HTDS evaluation or medical records with supporting documentation (Table IX.K-1). An additional 14 (0.4%) living evaluable participants had a diagnosis of hyperthyroidism based on medical records without supporting documentation, and 21 (0.6%) were based on a participant or his/her CATI respondent report. The cumulative incidence of hyperthyroidism was higher for women (9.0%) than men (2.3%).

It is important to understand that these 196 cases included a substantial number of iatrogenic cases (these are discussed below). Since endogenous hyperthyroidism was of particular importance, analyses that focused on cases of non-iatrogenic hyperthyroidism were given particular emphasis in this study.

Table IX.K-1. Basis for Diagnosis of Hyperthyroidism, by Sex

	Female		Male		Total	
Basis for Diagnosis	No.	%	No.	%	No.	%
Yes	157	9.0	39	2.3	196	5.7
▪ HTDS evaluation	77	4.4	18	1.1	95	2.8
▪ Medical records with supporting documentation	57	3.3	9	0.5	66	1.9
▪ Medical records *without* supporting documentation	12	0.7	2	0.1	14	0.4
▪ Participant/respondent report	11	0.6	10	0.6	21	0.6
No	1572	90.0	1649	97.4	3221	93.6
Unknown	18	1.0	5	0.3	23	0.7
Total	1747	100.0	1693	100.0	3440	100.0

Twenty-three living evaluable participants were classified as "unknown" with regard to diagnosis of hyperthyroidism. These participants did not have a medical record indicating hyperthyroidism, but 13 had a participant report of an unknown thyroid problem, with most indicating it was either an over or under active thyroid for which they took some type of medication. Eight others had no blood draw and for two others a diagnosis of hyperthyroidism could not be ruled out. These 23 participants were included as non-cases in analyses of the dose-response for hyperthyroidism.

One or more possible etiologies were identified for all of the 196 participants with hyperthyroidism. Exogenous thyroid medication was the most common etiology (59.2%) (Table IX.K-2). Graves disease (19.9%) was the second most frequent etiology of hyperthyroidism, followed by uncertain (14.3%). Six of the eight living evaluable participants with hyperthyroidism and an etiology of other were due to possible subacute thyroiditis, while two were due to possible Graves disease.

Table IX.K-2. Etiologies of Hyperthyroidism, by Sex

	Female		Male		Total	
Etiology	No.	%	No.	%	No.	%
Graves disease	32	20.4	7	17.9	39	19.9
Toxic nodular goiter	2	1.3	0	--	2	1.0
Solitary autonomous nodule	1	0.6	1	2.6	2	1.0
Subacute thyroiditis	3	1.9	2	5.1	5	2.6
Silent/post-partum thyroiditis	1	0.6	0	--	1	0.5
Exogenous thyroid medication	102	65.0	14	35.9	116	59.2
Uncertain	13	8.3	15	38.5	28	14.3
Other	8	5.1	0	--	8	4.1
Total with hyperthyroidism	157	100.0	39	100.0	196	100.0

Note: A participant can have >1 etiology

K.1.a. Non-iatrogenic Hyperthyroidism

Since the inclusion of iatrogenic hyperthyroidism might mask an effect of radiation on risk of endogenous hyperthyroidism, an additional disease outcome of non-iatrogenic hyperthyroidism was also defined. A total of 50 living evaluable participants had diagnoses of non-iatrogenic hyperthyroidism based on their HTDS evaluations or on medical records with supporting documentation (Table IX.K-3).

Table IX.K-3. Non-Iatrogenic Hyperthyroidism, by Sex

	Female		Male		Total	
Non-iatrogenic Hyperthyroidism	No.	%	No.	%	No.	%
Yes	37	2.1	13	0.8	50	1.4
No	1710	97.9	1680	99.2	3390	98.5
Total	1747	100.0	1693	100.0	3440	100.0

K.2. Analysis of Hyperthyroidism Risk

K.2.a. Primary Analysis

Of the 161 living evaluable participants with a diagnosis of hyperthyroidism based on the HTDS examination or medical records with supporting documentation, six were out-of-area participants for whom the CIDER program could not calculate dose estimates. The proportions with hyperthyroidism are shown by sex, dose category and basis for diagnosis in Table IX.K-4. The numbers and proportions with diagnoses of non-iatrogenic hyperthyroidism are also shown.

Table IX.K-4. Diagnoses of Hyperthyroidism by Sex, Estimated Dose, and Basis for Diagnosis

A. Female

Thyroid Radiation Dose (mGy)	Living Evaluable Female	Primary Definition: Cases Based on HTDS or Med. Rec. with Supporting Documentation		1st Alternative Definition: Cases based on HTDS or Med. Rec. with or without Supporting Documentation		2nd Alternative Definition: Cases Based on Any Diagnosis or Participant or CATI Report		Non-iatrogenic Hyperthyroidism (Primary Definition)	
	No.	No.	%	No.	%	No.	%	No.	%
Out of Area	125	6	4.8	6	4.8	7	5.6	2	1.6
< 10	182	10	5.5	12	6.6	13	7.1	5	2.7
10-49	320	27	8.4	29	9.1	29	9.1	6	1.9
50-99	313	30	9.6	32	10.2	36	11.5	9	2.9
100-149	220	9	4.1	10	4.5	12	5.5	2	0.9
150-199	126	11	8.7	14	11.1	16	12.7	0	0.0
200-299	139	13	9.4	15	10.8	15	10.8	4	2.9
300-399	144	14	9.7	14	9.7	15	10.4	4	2.8
400-999	171	13	7.6	13	7.6	13	7.6	5	2.9
1000+	7	1	14.3	1	14.3	1	14.3	0	--
Total	1747	134	7.7	146	8.4	157	9.0	37	2.1

Table IX.K-4. Diagnoses of Hyperthyroidism by Sex, Estimated Dose, and Basis for Diagnosis (continued)

B. Male

Thyroid Radiation Dose (mGy)	Living Evaluable Male	Primary Definition: Cases Based on HTDS or Med. Rec. with Supporting Documentation		1st Alternative Definition: Cases based on HTDS or Med. Rec. with or without Supporting Documentation		2nd Alternative Definition: Cases Based on Any Diagnosis or Participant or CATI Report		Non-iatrogenic Hyperthyroidism (Primary Definition)	
	No.	No.	%	No.	%	No.	%	No.	%
Out of Area	124	0	--	0	--	0	--	0	--
< 10	186	5	2.7	5	2.7	5	2.7	2	1.1
10-49	314	4	1.3	4	1.3	6	1.9	2	0.6
50-99	310	3	1.0	5	1.6	6	1.9	2	0.6
100-149	171	4	2.3	4	2.3	5	2.9	4	2.3
150-199	109	1	0.9	1	0.9	2	1.8	1	0.9
200-299	148	1	0.7	1	0.7	2	1.4	1	0.7
300-399	160	6	3.8	6	3.8	7	4.4	1	0.6
400-999	154	2	1.3	2	1.3	5	3.2	0	--
1000+	17	1	5.9	1	5.9	1	5.9	0	--
Total	1693	27	1.6	29	1.7	39	2.3	13	0.8

Parameter estimates for the linear dose-response model based on the 3191 in-area participants are shown in Table IX.K-5 below. Based on maximum likelihood analysis of the sex-stratified linear probability model, and using the primary dose estimates, the estimated slope B was 0.011 per Gy; row 1 of Table IX.K-5). The Bonferroni-adjusted 95% confidence interval ranged from less than –0.008 to 0.052, thus the cumulative incidence of hyperthyroidism did not increase significantly with increasing dose (p = 0.22). The corresponding estimated background rates for diagnosis of hyperthyroidism were 0.077 with confidence interval (0.060, 0.094) for women and 0.015 with confidence interval (0.006, 0.025) for men. Similar results were obtained when the model was fit by the method of least squares using ungrouped or grouped data (Table IX.K-5, rows 2 and 3).

Table IX.K-5. Summary of Dose-Response Results for Diagnoses of Hyperthyroidism

Row	Outcome	Dose Response Model	Dose Estimates	Exclusions Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
1.	Primary definition (HTDS evaluation or medical record with documentation)	Linear	Primary	None	MLE	.077 ± .007 (.060, .094)	.015 ± .004 (.006, .025)	.011 ± .015 (<-.008, .052)	0.22
2.	Primary definition	Linear	Primary	None	LSU	.076 ± .006 (.062, .090)	.014 ± .006 (0*, .029)	.018 ± .017 (-.022, .058)	0.15
3.	Primary definition	Linear	Primary	None	LSG	.077 ± .006 (.062, .092)	.015 ± .006 (0*, .030)	.012 ± .019 (-.034, .059)	0.26
4.	Alternative def. #1(HTDS or medical record with or without documentation)	Linear	Primary	None	MLE	.085 ± .007 (.067, .103)	.017 ± .004 (.007, .027)	.007 ± .015 (<-.008, .049)	0.32
5.	Alternative def. #2 (Any diagnosis or participant/respondent report)	Linear	Primary	None	MLE	.090 ± .008 (.071, .109)	.022 ± .005 (.011, .034)	.015 ± .018 (<-.011, .063)	0.19

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model. "0*" indicates that the lower confidence limit for a background rate was less than 0.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.K-5. Summary of Dose-Response Results for Diagnoses of Hyperthyroidism (continued)

Row	Outcomes	Dose-Response Model	Dose Estimates	Exclusions/ Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
6.	Non-iatrogenic hyperthyroidism	Linear	Primary	None	MLE	.022 ± .004 (.012, .033)	.009 ± .003 (.002, .015)	-.004 ± .013 (NE, .019)	0.78
7.	Primary definition	LQ	Primary	None	LSU	.077 ± .007 (.061, .093)	.015 ± .007 (0*, .032)	Lin: .009 ± .029 (–.062, .080) Quad: .007 ± .019 (–.040, .054)	Quad: 0.71
8.	Primary definition	Logistic	Primary	None	MLE	.074 (.058, .095)	.016 (.010, .026)	0.35 ± 0.32 (-.43, 1.12)	0.16
9.	Primary definition	Linear	Primary	Exclude dose > 1000 mGy	MLE	.078 ± .007 (.060, .096)	.016 ± .004 (.006, .027)	.003 ± .018 (<-.021, .051)	0.44
10.	Primary definition	Linear	Primary	Exclude dose > 400 mGy	MLE	.076 ± .008 (.057, .094)	.014 ± .004 (.004, .024)	.028 ± .027 (-.029, .101)	0.13
11.	Primary definition	Linear	Primary	Exclude Ok and F/S geostrata	MLE	.073 ± .007 (.055, .090)	.009 ± .004 (.0004, .018)	.025 ± .016 (<-.006, .067)	0.046

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model. "0*" indicates that the lower confidence limit for a background rate was less than 0.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.K-5. Summary of Dose-Response Results for Diagnoses of Hyperthyroidism (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions/ Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
12.	Primary definition	Linear	Alt. #1	None	MLE	.080 ± .007 (.062, .098)	.018 ± .004 (.008, .028)	-.005 ± .015 (<-.008, .034)	0.63
13.	Primary definition	Linear	Alt. #2	None	MLE	.079 ± .007 (.062, .097)	.018 ± .004 (.007, .028)	-.002 ± .015 (<-.008, .037)	0.55
14.	Primary definition	Linear	Primary	Include OOA (scoping analysis #1)	MLE	.074 ± .007 (.057, .090)	.013 ± .004 (.005, .022)	.017 ± .015 (<-.007,>.059)	0.11
15.	Primary definition	Linear	Primary	Include OOA (scoping analysis #2)	MLE	.074 ± .007 (.058, .090)	.013 ± .004 (.005, .022)	.016 ± .015 (<-.007,> .058)	0.12

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model. "0*" indicates that the lower confidence limit for a background rate was less than 0.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

K.2.b. Alternative Definitions for Diagnosis of Hyperthyroidism

Two alternative definitions for cases of hyperthyroidism were considered. The first alternative added 14 cases with diagnoses based on medical records without supporting documentation, for a total of 175 cases. The resulting dose-response had estimated slope of 0.007 per Gy with Bonferoni-adjusted confidence interval ranging from less than -0.008 to 0.049 per Gy (Table IX.K-5, row 4). The second alternative criterion for defining cases of hyperthyroidism added another 21 participants based solely on a report from the participant or his/her CATI respondent, for a total of 196 cases, estimated slope of 0.015 per Gy, and confidence interval ranging from less than – 0.011 to 0.063 per Gy (Table IX.K-5, row 5). The parameter estimates for the linear dose-response model [1] were not significantly greater than zero ($p = 0.32$ and $p = 0.19$), showing no evidence that the cumulative incidence of hyperthyroidism increased with increasing dose for either alternative criterion of hyperthyroidism.

K.2.b.1. Non-iatrogenic Hyperthyroidism

In the analyses described above, the participants with iatrogenic hyperthyroidism were included among the cases. In order to focus on endogenous outcomes, an additional analysis was performed in which participants with iatrogenic hyperthyroidism only were excluded from the cases. This left a total of 48 cases of non-iatrogenic hyperthyroidism based on the HTDS examination or medical records with supporting documentation among the 3191 in-area evaluable participants. As shown in row 6 of Table IX.K-5, the dose-response was slightly negative, with estimated slope – 0.004 per Gy and upper 95% confidence limit 0.019 per Gy ($p=0.78$).

K.2.c Alternative Dose-Response Functions

As shown in row 7 of Table IX.K-5, the estimated regression coefficient for the dose-squared term in the linear-quadratic dose-response model [5] was 0.007 with Bonferroni-adjusted 95% confidence interval ranging from –0.040 to 0.054. Thus the addition of a quadratic term did not significantly improve the fit of the model ($p = 0.71$).

In the analysis of hyperthyroidism based on the HTDS examination or medical records with supporting documentation, i.e., the primary criterion for defining cases with hyperthyroidism, the regression parameter for the effect of dose in the sex-stratified logistic regression model [2] was estimated as 0.35 with Bonferroni-adjusted 95% confidence interval ranging from –0.43 to 1.12 (Table IX.K-5, row 8). Thus the cumulative incidence of hyperthyroidism did not increase significantly with increasing dose ($p = 0.16$).

K.2.d. Effect of Excluding Participants in High Dose Categories

As shown in rows 9 and 10 of Table IX.K-5, when participants in high dose categories were excluded, the cumulative incidence of hyperthyroidism did not increase significantly with increasing dose ($p = 0.44$ and $p = 0.13$ when participants with estimated dose > 1000 mGy and > 400 mGy were excluded, respectively).

K.2.e. Effect of Excluding Okanogan and Ferry/Stevens Geostrata

When the Okanogan and Ferry/Stevens geostrata were excluded, the estimated slope of the sex-stratified linear dose-response model was 0.025 per Gy with Bonferroni-adjusted 95% confidence interval ranging from less than –0.006 to 0.067 per Gy (Table IX.K-5, row 11). While this result might be regarded as evidence that the cumulative incidence of hyperthyroidism increased with increasing dose

among the participants in the remaining geostrata (p = 0.046), it is not considered statistically significant in view of the large number of significance tests that were performed.

K.2.f. Analysis of Hyperthyroidism in Relation to Alternative Dose Estimates

As shown in rows 12 and 13 of Table IX.K-5, the slope of the dose-response was slightly negative when estimated in relation to either of the alternative dose estimates. Thus there was no evidence from these analyses that risk of hyperthyroidism increased significantly with increasing dose (p = 0.63 and p = 0.55 for the first and second alternative dose estimates, respectively).

K.2.g. Scoping Analysis Regarding Out-of-Area Participants

See section VIII.C.1.a.3 for a description of the scoping analyses that were performed to assess the possible impact of including the 249 out-of-area participants. As summarized in rows 14 and 15 of Table IX.K-5, in neither scoping analysis did the cumulative incidence of hyperthyroidism increase significantly with increasing dose (p = 0.11 and p = 0.12 for the first and second scoping analyses, respectively).

K.2.h. Analysis of Hyperthyroidism in Relation to Alternative Representations of Exposure

In the analyses by geostratum and by dichotomous exposure variable, the sex and age-adjusted comparisons of cumulative incidence were performed as described in section VIII.C.2.a.2.

K.2.h.1. Analysis by Geostratum

As shown in Table IX.P-18, among the 3429 living evaluable in area or out-of-area participants with ultrasound results, the proportions with palpable UDAs ranged from 9/68 (13.2% in the Stevens/Ferry Counties geostratum) to 9/177 (5.1%, Richland) for women, and from 5/63 (7.9%, Okanogan County) to 13/501 (2.6%, Pasco/Kennewick) for men (p = 0.051 for heterogeneity among the nine geostrata). In particular the percentages with palpable UDAs were somewhat higher in the Okanogan and Ferry/Stevens geostrata (12.6% for women, 6.8% for men) than in the remaining geostrata (8.5% and 3.9%, respectively; p = 0.0086). Since it was likely that participants in the Okanogan and Ferry/Stevens geostrata tended to have lower thyroid doses from Hanford's ^{131}I than those in other geostrata, it does not appear that these differences can be attributed to an effect of Hanford's ^{131}I.

Table IX.K-6. Diagnoses of Hyperthyroidism Based on the HTDS Evaluation or on Medical Records with Supporting Documentation, by Geostratum and Sex

	Female			Male			Total		
Geostratum	No.	Cases	%	No.	Cases	%	No.	Cases	%
Richland	179	13	7.3	173	2	1.2	352	15	4.3
Pasco/Kennewick	508	37	7.3	501	6	1.2	1009	43	4.3
Benton County	376	30	8.0	358	4	1.1	734	34	4.6
Franklin County	73	4	5.5	76	1	1.3	149	5	3.4
Adams County	165	14	8.5	156	2	1.3	321	16	5.0
Walla Walla (city)	133	7	5.3	131	4	3.1	264	11	4.2
Walla Walla County	170	14	8.2	164	1	0.6	334	15	4.5
Okanogan County	75	8	10.7	64	3	4.7	139	11	7.9
Ferry/Stevens Counties	68	7	10.3	70	4	5.7	138	11	8.0
Total	1747	134	7.7	1693	27	1.6	3440	161	4.7

K.2.h.2. Analysis by Dichotomous Exposure Variable

Fifty-six (4.5%) of the 1257 participants included in these analyses had a diagnosis of hyperthyroidism based on the HTDS examination or medical records with supporting documentation (see Table IX.K-7). These included 28/580 (4.8%) in the high exposure group and 28/677 (4.1%) in the low exposure group. The cumulative incidence of hyperthyroidism was not significantly higher in the high exposure group (p = 0.074).

Table IX.K-7. Diagnoses of Hyperthyroidism Based on HTDS or Medical Record with Supporting Documentation, by Exposure Group and Sex

	Female			Male			Total		
Exposure Group	No.	Cases	%	No.	Cases	%	No.	Cases	%
Low	351	22	6.3	326	6	1.8	677	28	4.1
High	298	24	8.1	282	4	1.4	580	28	4.8
Total	649	46	7.1	608	10	1.6	1257	56	4.5

The second alternative criterion for defining cases of hyperthyroidism included all possible cases (see section IX.K.1 above). Among the 1257 participants included in these analyses, use of the second alternative added three cases based on medical records without supporting documentation and five others based on a report from the participant or his/her CATI respondent, for a total of 64 (5.1%). Since the number of added cases was small, the logistic regression analysis with adjustment for the effects of sex and age at HTDS examination gave essentially the same results as that based on the primary definition for hyperthyroidism. In particular, the cumulative incidence of hyperthyroidism was not significantly higher in the high exposure group (p = 0.062).

K.2.i. Confounding and Effect Modification

As described in section VIII above, additional sex-stratified logistic regression models were investigated to examine the possibility that the primary dose-response results might be influenced by confounding, and to search for factors that might modify a radiation dose-response. These analyses were based on the primary definition of hyperthyroidism; those based on an HTDS diagnosis or on medical records with documented diagnoses, and on the primary dose estimates. Table IX.K-8 displays results for models including sex, age at first exposure to Hanford ^{131}I (prenatal, or < 180 days), age at HTDS examination, estimated dose from the NTS, history of any cancer other than thyroid, and HTDS interview type.

Note that sex was not analyzed as a possible confounder since its effect was already adjusted for in the sex-stratified model. None of the other factors in Table IX.K-8 appears to be a confounder: for none does the adjusted estimate of the regression coefficient differ markedly from the unadjusted estimate. Therefore, it does not appear that omitting these factors introduces any important bias in the dose-response results.

The analyses of effect modification address the question of whether the dose-response might vary according to other characteristics of the study participants. This was tested by comparing the estimated regression coefficients for the groups defined by each covariate. As shown in Table IX.K-8, the regression coefficients did not differ significantly between the groups defined by any of the covariates, with the possible exception of estimated NTS dose. The estimated regression coefficient was 0.828 with confidence interval (–0.174, 1.83) for the 1622 participants with estimated NTS dose ≤ 5.3 mGy, compared to –1.18 with confidence interval (–3.58, 1.22) for those with higher estimated NTS doses. The p-value for comparing these two slopes, 0.019, should be interpreted with caution in view of the large number of significance tests that were performed, and of the extensive overlap of the two estimates' confidence intervals.

Table IX.K-8. Confounding and Effect Modification by Sex, Age at Exposure or HTDS Examination, Estimated Thyroid Dose from NTS, History of Any Cancer Other than Thyroid, and Interview Type: Hyperthyroidism

		Estimated Dose-Response Coefficient (per Gy)				
Covariate (0=No, 1=Yes)	Yes / Total	Unadjusted	Adjusted for Confounding	Including Effect Modification Group 0	Including Effect Modification Group 1	P
Female?	1622 / 3191	.346 ± .325 (–.432, 1.12)	Not Applicable	.282 ± .760 (–1.62, 2.18)	.361 ± .360 (–.538, 1.26)	.92
Prenatal exposure?	1034 / 3191	.346 ± .325 (–.432, 1.12)	.333 ± .328 (–.512, 1.18)	.439 ± .356 (–.501, 1.38)	-.089 ± .774 (–2.13, 1.95)	.52
1st exposure before age 180 days?	1478 / 3191	.346 ± .325 (–.432, 1.12)	.355 ± .329 (–.491, 1.20)	.829 ± .582 (–.706, 2.36)	.136 ± .435 (–1.01, 1.28)	.34
Age at exam > 50?	2001 / 3191	.346 ± .325 (–.432, 1.12)	.392 ± .323 (–.441, 1.23)	.249 ± .564 (–1.24, 1.74)	.471 ± .397 (–.577, 1.52)	.74
NTS ^{131}I dose > 5.3 mGy?	1567 / 3189	.346 ± .325 (–.432, 1.12)	.314 ± .334 (–.547, 1.18)	.828 ± .380 (–.174, 1.83)	–1.18 ± .909 (–3.58, 1.22)	.019
History of any cancer other than thyroid cancer?	248 / 3186	.346 ± .325 (–.432, 1.12)	.384 ± .331 (–.469, 1.24)	.489 ± .345 (–.421, 1.40)	–1.43 ± 2.51 (–8.04, 5.19)	.30
Expanded In-Person Interview?	1212 / 3191	.346 ± .325 (–.432, 1.12)	.463 ± .327 (–.380, 1.31)	.520 ± .547 (–.924, 1.96)	.431 ± .413 (–.659, 1.52)	.90

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

Tables IX.K-9 and IX.K-10 display similar results from analyses including history of medical or dental x-ray exposure or of occupational exposures as potential confounding or effect modifying factors. The estimates of the regression coefficient calculated with adjustment for confounding are all close to the unadjusted estimates. Thus there was no evidence that a confounding effect of any of these covariates has obscured a positive dose-response for hyperthyroidism.

There is no evidence of any statistically significant effect modification by any of the covariates in Tables IX.K-9 and IX.K-10, with the possible exception of history of diagnostic x-rays of chest or upper body, including mammograms (p = 0.031; Table IX.K-9). The estimated dose-response coefficient was markedly negative (–7.82) for the 352 participants without such histories, and not markedly different from zero for the majority of participants (0.432 with confidence interval ranging from –0.407 to 1.27). The statistical significance of this difference must be interpreted with caution due to the large number of such comparisons that were performed. Moreover the difference consists of a very negative dose-response in a minority of participants. Therefore it does not appear that any of the covariates in Tables IX.K-9 and IX.K-10 identified a group in which a clearly significant dose-response was present.

Table IX.K-9. Confounding and Effect Modification by History of Medical and Dental Radiation Exposures: Hyperthyroidism

Have You Ever Had: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy) Unadjusted	Adjusted for Confounding	Including Effect Modification Group 0	Group 1	P
CAT scan of the upper body?	775 / 3149	.347 ± .325 (–.430, 1.13)	.349 ± .329 (–.497, 1.20)	.400 ± .354 (–.534, 1.33)	.107± .797 (–2.00, 2.21)	.74
Diagnostic x-rays of the head?	1191 / 3155	.356 ± .324 (–.419, 1.13)	.367 ± .323 (–.467, 1.20)	.149 ± .452 (–1.04, 1.34)	.667 ± .480 (–.598, 1.93)	.43
Diagnostic x-rays of the neck?	966 / 3167	.345 ± .327 (–.439, 1.13)	.348 ± .329 (–.500, 1.20)	.190 ± .511 (–1.16, 1.54)	.464 ± .417 (–.637, 1.57)	.68
Diagnostic x-rays of chest or upper body, including mammograms?	2821 / 3173	.345 ± .325 (–.433, 1.12)	.347 ± .325 (–.490, 1.18)	–7.82 ± 5.01 (–21.0, 5.41)	.432 ± .318 (–.407, 1.27)	.031
Diagnostic x-rays of the stomach or mid-back?	692 / 3120	.310 ± .334 (–.489, 1.11)	.302 ± .334 (–.559, 1.16)	.107 ± .401 (–.952, 1.17)	.945 ± .639 (–.741, 2.63)	.28
Barium enema?	825 / 3159	.330 ± .327 (–.454, 1.11)	.332 ± .328 (–.512, 1.18)	.586 ± .360 (–.364, 1.53)	–.478 ± .785 (–2.55, 1.59)	.19
Upper GI?	1146 / 3177	.348 ± .324 (–.428, 1.12)	.343 ± .325 (–.494, 1.18)	.541 ± .381 (–.464, 1.55)	–.054 ± .606 (–1.65, 1.54)	.39
Intravenous pyelogram?	398 / 3157	.322 ± .330 (–.468, 1.11)	.322 ± .332 (–.533, 1.18)	.374 ± .346 (–.539, 1.29)	–.082 ± 1.01 (–2.74, 2.57)	.66
Fluoroscopy of the upper body?	246 / 3161	.329 ± .329 (–.458, 1.12)	.338 ± .330 (–.512, 1.19)	.382 ± .336 (–.506, 1.27)	–.250 ± 1.30 (–3.68, 3.18)	.63
Nuclear scan (excluding thyroid scan)?	217 / 3162	.379 ± .322 (–.391, 1.15)	.378 ± .321 (–.449, 1.20)	.399 ± .323 (–.452, 1.25)	–.481 ± 2.20 (–6.29, 5.33)	.68
Dental x-rays that did not usually include a lead shield over the neck area?	1648 / 3191	.346 ± .325 (–.432, 1.12)	.342 ± .327 (–.499, 1.18)	.561 ± .422 (–.554, 1.68)	.055 ± .539 (–1.37, 1.48)	.45

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

Table IX.K-10. Confounding and Effect Modification by Occupational History: Hyperthyroidism

Have You Ever Worked in Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted	Adjusted for Confounding	Including Effect Modification		P
				Group 0	Group 1	
Any metal industry?	238 / 3191	.346 ± .325 (–.432, 1.12)	.351 ± .325 (–.485, 1.19)	.329 ± .332 (–.546, 1.20)	1.12 ± 1.85 (–3.76, 6.00)	.68
Any nuclear facility?	371 / 3191	.346 ± .325 (–.432, 1.12)	.372 ± .326 (–.468, 1.21)	.372 ± .341 (–.528, 1.27)	.366 ± 1.10 (–2.54, 3.27)	1.00
Any other industry or occupation where you may have been exposed to radioactive materials or x-rays?	442 / 3191	.346 ± .325 (–.432, 1.12)	.355 ± .325 (–.481, 1.19)	.265 ± .372 (–.716, 1.25)	.681 ± .648 (–1.03, 2.39)	.59
Any of the above industries or occupations?	892 / 3191	.346 ± .325 (–.432, 1.12)	.352 ± .325 (–.486, 1.19)	.320 ± .389 (–.706, 1.35)	.430 ± .592 (–1.13, 1.99)	.88

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

Table IX.K-11 displays the results of analyses of possible confounding or effect modification by smoking variables. There was no evidence that the dose-response was significantly confounded by either smoking variable, or that there was a dose-response that differed significantly according to smoking history.

Table IX.K-11. Confounding and Effect Modification by Smoking: Hyperthyroidism

Have You Ever Smoked Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted	Adjusted for Confounding	Including Effect Modification		P
				Group 0	Group 1	
Cigarettes (unfiltered or filtered)?	1854 / 3183	.344 ± .326 (–.436, 1.12)	.333 ± .325 (–.505, 1.17)	.296 ± .613 (–1.32, 1.91)	.347 ± .382 (–.661, 1.35)	.94
Any of cigarettes, cigar or pipe?	1900 / 3183	.344 ± .326 (–.436, 1.12)	.333 ± .325 (–.505, 1.17)	.294 ± .612 (–1.32, 1.91)	.349 ± .383 (–.661, 1.36)	.94

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

K.2.j. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for hyperthyroidism are shown in Figure IX.K-1 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in that figure are calculated at the 98.33% confidence level, i.e., are adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates. While the point estimate of the slope is greater than 0 for 81 of the 100 realizations, the confidence interval includes 0 for all 100 realizations. Also shown in Figure IX.K-1 (to the right of realization 100) are the

estimates and confidence intervals calculated using (from left to right) the median, geometric mean and mean of each participant's 100 dose realizations. In summary, for none of the 100 realizations of the estimated doses was there a statistically significant dose-response, although the estimated slope was greater than 0 for most realizations.

Figure IX.K-1. Plot of estimated Slope and 95% CI by Dose Realization: Hyperthyroidism

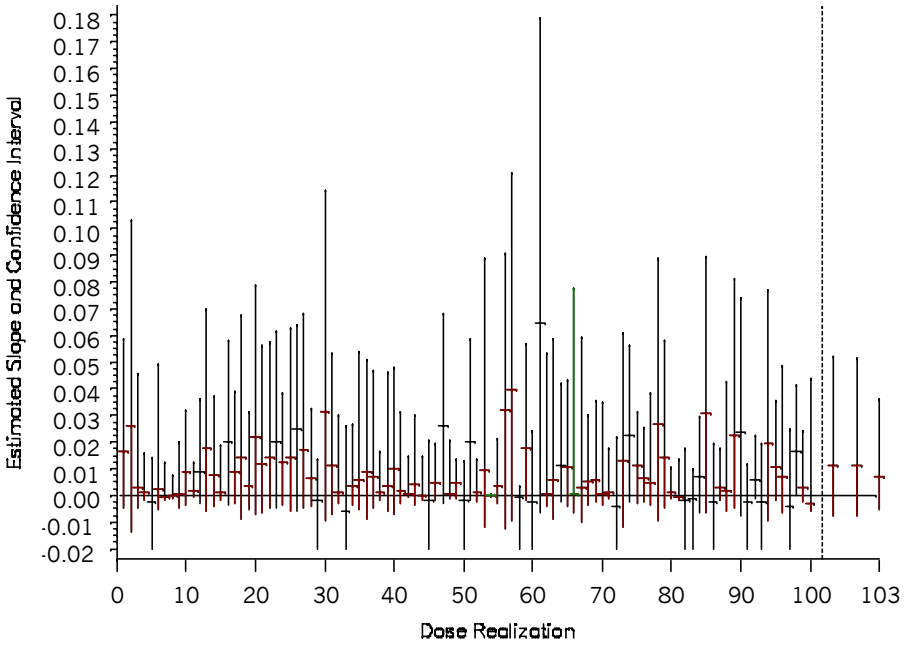
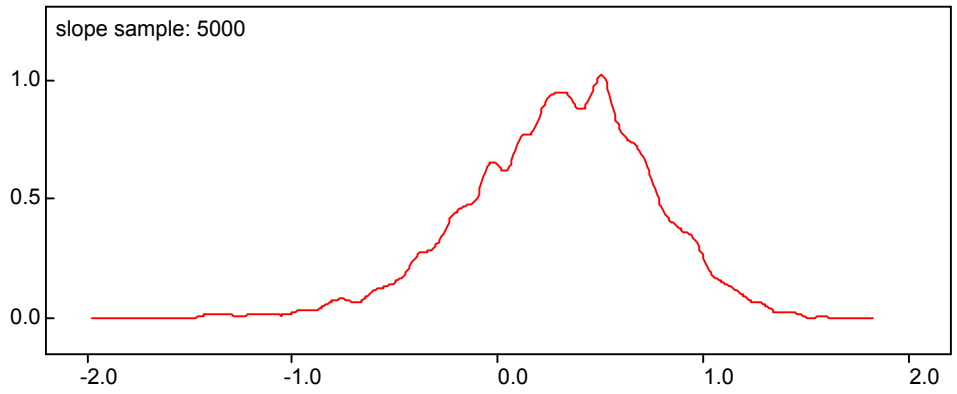


Figure IX.K-2 displays the distribution of the 5000 logistic regression coefficient estimates obtained by the simulation procedures described in section VIII.C.2.b.3 above. It is evident from the figure that most of the estimates were between about –1.0 and 1.5. The estimate was less than or equal to 0 for 1193 of the 5000 replications, implying an empirical one-tailed p-value of 0.24. The median estimate was 0.33, and the upper and lower percentiles corresponding to the Bonferroni-adjusted 95% confidence interval were –0.97 and 1.27. These may be compared to the estimates of 0.35 with the confidence interval (–0.43, 1.12) obtained using the median dose estimate without adjustment for uncertainty. Thus, this method of adjusting the estimated logistic regression coefficient for the uncertainty in the dose estimates did not provide evidence that the cumulative incidence of hypothyroidism increased with increasing dose.

Figure IX.K-2. Distribution of Simulation Estimates of Logistic Regression Coefficient: Hyperthyroidism



L. Multinodular Thyroid Gland

L.1. Occurrence of Multinodular Thyroid Gland

The primary and alternative definitions for multinodular thyroid gland were as follows:

- Primary definition: HTDS evaluation (95 cases)
- Alternative definition #1: HTDS evaluation or medical records (114 cases)
- Alternative definition #2: Any diagnosis or participant/respondent report (115 cases).

Of the 3440 living evaluable participants, 95 (2.8%) had a diagnosis of multinodular thyroid gland based on the HTDS evaluation (Table IX.L-1). An additional 19 (0.6%) living evaluable participants had a diagnosis of multinodular thyroid gland based on medical records, and 1 had a participant/respondent report of multinodular thyroid gland.

Table IX.L-1. Basis for Diagnosis of Multinodular Thyroid Gland, by Sex

	Female		Male		Total	
Basis for Diagnosis	No.	%	No.	%	No.	%
Yes	90	5.2	25	1.5	115	3.3
▪ HTDS evaluation	73	4.2	22	1.3	95	2.8
▪ Medical record	16	0.9	3	0.2	19	0.6
▪ Participant/respondent report	1	0.1	0	--	1	0.0
No	1656	94.8	1668	98.5	3324	96.6
Unknown	1	0.1	000	--	1	0.0
Total	1747	100.0	1693	100.0	3440	100.0

One living evaluable participant was classified as "unknown" with regard to diagnosis of multinodular thyroid gland. This participant did not have medical record or participant/respondent report indicating multinodular thyroid gland, and did not have an HTDS evaluation due to physician disagreement on the diagnosis, with one physician assigning a diagnosis of multinodular gland while the other assigned a diagnosis of solitary nodule. This participant was included as a non-case in analyses of the dose-response for multinodular thyroid gland.

As shown in Table IX.L-2, the most common etiology of multinodular thyroid gland was other (65.2%), followed by Hashimoto's thyroiditis (32.2%), and hypothyroidism (13.0%).

Table IX.L-2. Etiologies of Multinodular Thyroid Gland, by Sex

	Female		Male		Total	
Etiology	No.	%	No.	%	No.	%
Graves disease	3	3.3	0	--	3	2.6
Hashimoto's thyroiditis	30	33.3	7	28.0	37	32.2
Hypothyroidism	14	15.6	1	4.0	15	13.0
Other	56	62.2	19	76.0	75	65.2
Total with multinodular thyroid gland	90	100.0	25	100.0	115	100.0

Note: A participant can/may have more than one etiology

Of the 75 with an other etiology for multinodular thyroid gland, 47 (62.7%) were unknown/uncertain, 11 (14.7%) were colloid nodule, and 10 (13.3%) were probable/possible Hashimoto's thyroiditis (Table IX.L-3).

Table IX.L-3. "Other" Etiologies of Multinodular Thyroid Gland, by Sex

	Female		Male		Total	
Other Etiologies	No.	%	No.	%	No.	%
Unknown/uncertain	37	66.1	10	52.6	47	62.7
Probable/possible Hashimoto's thyroiditis	7	12.5	3	15.8	10	13.3
Colloid nodules	6	10.7	5	26.3	11	14.7
Colloid goiter	1	1.8	1	5.3	2	2.7
Papillary/follicular cancer	2	3.6	0	--	2	2.7
Possible hypothyroidism	1	1.8	0	--	1	1.3
Probable medical radiation	1	1.8	0	--	1	1.3
Multiple etiologies*	1	1.8	0	--	1	1.3
Total with an other etiology of multinodular thyroid gland	56	100.0	19	100.0	75	100.0

* Includes: 1) adenomatous nodules, 2) focus of papillary cancer, 3) focus of Hashimoto's

L.2. Analysis of Multinodular Thyroid Gland Risk

L.2.a. Primary Analysis

Of the 95 living evaluable participants with a diagnosis of multinodular thyroid gland based on the HTDS examination, 10 were out-of-area participants for whom the CIDER program could not calculate dose estimates. The proportions with multinodular thyroid gland are shown by sex, dose category and basis for diagnosis in Table IX.L-4.

Table IX.L-4. Diagnoses of Multinodular Thyroid Gland by Sex, Estimated Dose, and Basis for Diagnosis

A. Female

Thyroid Radiation Dose (mGy)	Living Evaluable Female	Primary Definition: Cases Based on HTDS Examination		1st Alternative Definition Cases Based on HTDS Examination or Medical Records		2nd Alternative Definition: Cases Based on Any Diagnosis or Participant or CATI Report	
	No.	No.	%	No.	%	No.	%
Out of Area	125	9	7.2	9	7.2	9	7.2
< 10	182	10	5.5	10	5.5	10	5.5
10-49	320	13	4.1	16	5.0	16	5.0
50-99	313	12	3.8	17	5.4	17	5.4
100-149	220	10	4.5	12	5.5	12	5.5
150-199	126	3	2.4	3	2.4	3	2.4
200-299	139	3	2.2	7	5.0	8	5.8
300-399	144	6	4.2	6	4.2	6	4.2
400-999	171	7	4.1	9	5.3	9	5.3
1000+	7	0	--	0	--	0	--
Total	1747	73	4.2	89	5.1	90	5.2

Table IX.L-4. Diagnoses of Multinodular Thyroid Gland by Sex, Estimated Dose, and Basis for Diagnosis (continued)

B. Male

Thyroid Radiation Dose (mGy)	Living Evaluable Male	Primary Definition: Cases Based on HTDS Examination		1st Alternative Definition Cases Based on HTDS Examination or Medical Records		2nd Alternative Definition: Cases Based on Any Diagnosis or Participant or CATI Report	
	No.	No.	%	No.	%	No.	%
Out of Area	124	1	0.8	2	1.6	2	1.6
< 10	186	4	2.2	4	2.2	4	2.2
10-49	314	5	1.6	5	1.6	5	1.6
50-99	310	4	1.3	4	1.3	4	1.3
100-149	171	2	1.2	3	1.8	3	1.8
150-199	109	1	0.9	1	0.9	1	0.9
200-299	148	2	1.4	2	1.4	2	1.4
300-399	160	2	1.3	2	1.3	2	1.3
400-999	154	1	0.6	2	1.3	2	1.3
1000+	17	0	--	0	--	0	--
Total	1693	22	1.3	25	1.5	25	1.5

Parameter estimates for the linear dose-response model based on the 3191 in-area participants are shown in Table IX.L-5 below. Based on maximum likelihood analysis of the sex-stratified linear probability model, and using the primary dose estimates, the estimated slope B was –0.006 per Gy; row 1 of Table IX.L-5). The lower limit of the Bonferroni-adjusted 95% confidence interval was not estimated, but the upper limit was 0.014 per Gy, and the cumulative incidence of multinodular thyroid gland did not increase significantly with increasing dose (p = 0.88). The corresponding estimated background rates for diagnosis of multinodular thyroid gland were 0.040 with confidence interval (0.027, 0.053) for women and 0.014 with confidence interval (0.006, 0.023) for men. When the model was fit by the method of least squares, the estimated slope using either ungrouped or grouped data was even more negative than the maximum likelihood estimate (Table IX.L-5, rows 2 and 3), thereby providing no evidence that risk of multinodular thyroid gland increased with increasing dose (p = 0.89 and 0.83).

Table IX.L-5. Summary of Dose-Response Results for Diagnoses of Multinodular Thyroid Gland

Row	Outcome	Dose Response Model	Dose Estimates	Exclusions/ Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
1.	Primary definition (HTDS evaluation)	Linear	Primary	None	MLE	.040 ± .005 (.027, .053)	.014 ± .004 (.006, .023)	-.006 ± .016 (NE, .014)	0.88
2.	Primary definition	Linear	Primary	None	LSU	.042 ± .005 (.031, .053)	.016 ± .005 (.005, .027)	-.016 ± .013 (-.046, .015)	0.89
3.	Primary definition	Linear	Primary	None	LSG	.042 ± .005 (.031, .053)	.016 ± .005 (.004, .027)	-.014 ± .015 (-.049, .021)	0.83
4.	Alternative def. #1 (HTDS or medical record)	Linear	Primary	None	MLE	.050 ± .006 (.037, .065)	.016 ± .004 (.006, .025)	-.006 ± .016 (NE, .018)	0.86
5.	Alternative def. #2 (Any diagnosis or participant/respondent report)	Linear	Primary	None	MLE	.051 ± .006 (.037, .065)	.016 ± .004 (.006, .025)	-.006 ± .016 (NE, .018)	0.86

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.L-5. Summary of Dose-Response Results for Diagnoses of Multinodular Thyroid Gland (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions/ Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
6.	Primary definition	LQ	Primary	None	LSU	.043 ± .005 (.031, .055)	.017 ± .005 (.004, .029)	Lin: –.021 ± .022 (–.075, .032) Quad: .005 ± .014 (–.031, .040)	Quad: 0.75
7.	Primary definition	Logistic	Primary	None	MLE	.045 (.031, .064)	.015 (.009, .027)	–.82 ± .65 (-2.37, .72)	0.92
8.	Primary definition	Linear	Primary	Exclude dose > 1000 mGy	MLE	.042 ± .005 (.029, .055)	.016 ± .004 (.007, .025)	-.016 ± .014 (NE, >.016)	0.92
9.	Primary definition	Linear	Primary	Exclude dose > 400 mGy	MLE	.042 ± .006 (.028, .056)	.017 ± .004 (.007, .028)	-.025 ± .023 (<-.059, .035)	0.86
10.	Primary definition	Linear	Primary	Exclude Ok and F/S geostrata	MLE	.039 ± .006 (.026, .053)	.015 ± .004 (.005, .024)	-.006 ± .016 (NE, .015)	0.88
11.	Primary definition	Linear	Alt. #1	None	MLE	.040 ± .005 (.027, .053)	.014 ± .004 (.004, .024)	-.006 ± .018 (NE, .010)	0.92

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.L-5. Summary of Dose-Response Results for Diagnoses of Multinodular Thyroid Gland (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions/ Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
12.	Primary definition	Linear	Alt. #2	None	MLE	.040 ± .005 (.028, .053)	.014 ± .004 (.006, .023)	-.006 ± .010 (NE, .015)	0.86
13.	Primary definition	Linear	Primary	Include OOA (scoping analysis #1)	MLE	.043 ± .005 (.030, .055)	.014 ± .003 (.006, .022)	-.006 ± .015 (NE, .014)	0.88
14.	Primary definition	Linear	Primary	Include OOA (scoping analysis #2)	MLE	.043 ± .005 (.030, .055)	.014 ± .003 (.006, .022)	-.006 ± .015 (NE, >.013)	0.88

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

L.2.b. Alternative Definitions for Diagnosis of Multinodular Thyroid Gland

Two alternative definitions for cases of multinodular thyroid gland were considered. The first alternative added 19 cases with diagnoses based on medical records, for a total of 114 cases. The second added a single case based solely on a report from the participant or his/her CATI respondent, for a total of 115 cases. This last case had an estimated dose of 254 mGy.

As shown in rows 4 and 5 of Table IX.L-5, the slope of the dose-response in the sex-stratified linear model estimated for both alternative definitions of multinodular thyroid gland were nearly identical to the estimate based on the primary definition (estimated slope –0.006 per Gy with Bonferoni-adjusted 95% upper confidence limit 0.018 per Gy). Thus for neither alternative definition did the cumulative incidence of multinodular thyroid gland increase significantly with increasing dose ($p = 0.86$ for both alternatives).

L.2.c. Alternative Dose-Response Functions

As shown in row 6 of Table IX.L-5, the estimated regression coefficient for the dose-squared term in the linear-quadratic dose-response model [5] was 0.005 with Bonferroni-adjusted 95% confidence interval ranging from –0.031 to 0.040. Thus the addition of a quadratic term did not significantly improve the fit of the model ($p = 0.75$).

In the analysis of multinodular thyroid gland based on the HTDS examination, i.e., the primary criterion for defining cases with multinodular thyroid gland, the regression parameter for the effect of dose in the sex-stratified logistic regression model [2] was estimated as –0.82 with Bonferroni-adjusted 95% confidence interval ranging from -2.37 to 0.72 (Table IX.L-5, row 7). Thus the cumulative incidence of multinodular thyroid gland did not increase significantly with increasing dose ($p = 0.92$).

L.2.d. Effect of Excluding Participants in High Dose Categories

Rows 8 and 9 of Table IX.L-5 show the effect of excluding patients in high dose categories from the analysis of the sex-stratified linear dose-response model. When participants with estimated dose > 1000 mGy were excluded from the analysis, the estimated slope of the dose-response decreased to –0.016 per Gy, providing no evidence that the cumulative incidence of multinodular gland increased with increasing dose ($p = 0.92$). Similarly, when participants with estimated dose > 400 mGy were excluded, the estimated slope B was even more negative (–0.025, with Bonferroni-adjusted 95% confidence interval ranging from less than –0.059 to 0.035), providing no evidence that the cumulative incidence increased with increasing dose ($p = 0.86$).

L.2.e. Effect of Excluding Okanogan and Ferry/Stevens Geostrata

When participants in the Ferry/Stevens geostrata were excluded, the estimated slope B was virtually unchanged (–0.006 per Gy, Table IX.L-5, row 10), providing no evidence that the cumulative incidence of multinodular thyroid gland increased with increasing dose ($p = 0.88$).

L.2.f. Analysis of Multinodular Thyroid Gland in Relation to Alternative Dose Estimates

Rows 11 and 12 of Table IX.L-5 show that the estimated dose-response was almost unchanged when the two alternative sets of dose estimates were used in place of the primary doses. In particular there

was no evidence that cumulative incidence of multinodular thyroid gland increased with increasing doses from either of the alternative dose sets (p = 0.92 and p= 0.86 for alternative dose sets 1 and 2, respectively).

L.2.g. Scoping Analysis Regarding Out-of-Area Participants

See section VIII.C.1.a.3 for a description of the scoping analyses that were performed to assess the possible impact of including the 249 out-of-area participants. As summarized in rows 13 and 14 of Table IX.L-5, in neither scoping analysis did the cumulative incidence of multinodular thyroid gland increase significantly with increasing dose (p = 0.88 for both scoping analyses).

L.2.h. Analysis of Multinodular Thyroid Gland in Relation to Alternative Representations of Exposure

In the analyses by geostratum and by dichotomous exposure variable, the sex and age-adjusted comparisons of cumulative incidence were performed as described in section VIII.C.2.a.2.

L.2.h.1. Analysis by Geostratum

As shown in Table IX.L-6, the proportions of participants with multinodular thyroid gland (from the HTDS examination) ranged from 6/75 (8.0%) in the Okanogan geostratum to 0/179 (0%) in the Richland geostratum for women, and from 5/156 (3.2%) in the Adams geostratum to 0/70 (0%) in the Ferry/Stevens geostratum for men. The heterogeneity among the nine geostrata was not statistically significant (p=0.058). The proportions with multinodular thyroid gland were somewhat higher for women in the Okanogan and Ferry/Stevens geostrata (7.0%) compared to the other strata (3.9%), but this was not the case for men (0.7% for Okanogan and Ferry/Stevens versus 1.3% for the others, p = 0.29).

Table IX.L-6. Diagnoses of Multinodular Thyroid Gland Based on the HTDS Evaluation, by Geostratum and Sex

	Female			Male			Total		
Geostratum	No.	Cases	%	No.	Cases	%	No.	Cases	%
Richland	179	0	--	173	4	2.3	352	4	1.1
Pasco/Kennewick	508	17	3.3	501	2	0.4	1009	19	1.9
Benton County	376	16	4.3	358	3	0.8	734	19	2.6
Franklin County	73	2	2.7	76	2	2.6	149	4	2.7
Adams County	165	8	4.8	156	5	3.2	321	13	4.0
Walla Walla (city)	133	10	7.5	131	1	0.8	264	11	4.2
Walla Walla County	170	10	5.9	164	4	2.4	334	14	4.2
Okanogan County	75	6	8.0	64	1	1.6	139	7	5.0
Ferry/Stevens Counties	68	4	5.9	70	0	--	138	4	2.9
Total	1747	73	4.2	1693	22	1.3	3440	95	2.8

L.2.h.2. Analysis by Dichotomous Exposure Variable

Forty-six (3.7%) of the 1257 participants included in these analyses had a diagnosis of multinodular thyroid gland based on the HTDS examination (see Table IX.L-7). These included 19/580 (3.3%) in the high exposure group and 27/677 (4.0%) in the low exposure group. Thus the cumulative incidence of multinodular thyroid gland was not significantly higher in the high exposure group (p = 0.74).

Table IX.L-7. Diagnoses of Multinodular Thyroid Gland Based on HTDS Examination, by Exposure Group and Sex

	Female			Male			Total		
Exposure Group	No.	Cases	%	No.	Cases	%	No.	Cases	%
Low	351	23	6.6	326	4	1.2	677	27	4.0
High	298	16	5.4	282	3	1.1	580	19	3.3
Total	649	39	6.0	608	7	1.2	1257	46	3.7

L.2.i. Confounding and Effect Modification

As described in section VIII above, additional sex-stratified logistic regression models were investigated to examine the possibility that the primary dose-response results might be influenced by confounding, and to search for factors that might modify a radiation dose-response. These analyses were based on the primary definition of multinodular thyroid gland (HTDS diagnosis), and on the primary dose estimates. Table IX.L-8 displays results for models including sex, age at first exposure to Hanford ^{131}I (prenatal, or < 180 days), age at HTDS examination, estimated dose from the NTS, history of any cancer other than thyroid, and HTDS interview type.

Note that sex was not analyzed as a possible confounder since its effect was already adjusted for in the sex-stratified model. None of the other factors in Table IX.L-8 appears to be a confounder: for none does the adjusted estimate of the regression coefficient differ markedly from the unadjusted estimate. Therefore, it does not appear that omitting these factors introduces any important bias in the dose-response results.

The analyses of effect modification address the question of whether the dose-response might vary according to other characteristics of the study participants. This was tested by comparing the estimated regression coefficients for the groups defined by each covariate. As shown in Table IX.L-8, the regression coefficients did not differ significantly between the groups defined by any of the covariates. For two covariates the effect modification approached statistical significance: age at first exposure to Hanford's ^{131}I within the HEDR domain ($p = 0.061$), and estimated NTS dose ($p = 0.053$). However, for neither covariate was there evidence of a significant dose-response within a subgroup of participants. In fact for both of these covariates, the difference was due primarily to a very negative regression coefficient in one group.

Table IX.L-8. Confounding and Effect Modification by Sex, Age at Exposure or HTDS Examination, Estimated Thyroid Dose from NTS, History of Any Cancer Other than Thyroid, and Interview Type: Multinodular Thyroid Gland

		Estimated Dose-Response Coefficient (per Gy)				
Covariate (0=No, 1=Yes)	Yes / Total	Unadjusted	Adjusted for Confounding	Including Effect Modification Group 0	Including Effect Modification Group 1	P
Female?	1622 / 3191	−.824 ± .647 (−2.37, .724)	Not Applicable	−1.97 ± 1.57 (−5.89, 1.94)	−.513 ± .698 (−2.26, 1.23)	.37
Prenatal exposure?	1034 / 3191	−.824 ± .647 (−2.37, .724)	−.862 ± .656 (−2.55, .828)	−.740 ± .751 (−2.72, 1.24)	−1.21 ± 1.33 (−4.72, 2.30)	.76
1st exposure before age 180 days?	1478 / 3191	−.824 ± .647 (−2.37, .724)	−.840 ± .653 (−2.52, .843)	.319 ± .824 (−1.86, 2.49)	−2.42 ± 1.25 (−5.72, .874)	.052
Age at exam > 50?	2001 / 3191	−.824 ± .647 (−2.37, .724)	−.672 ± .641 (−2.32, .978)	−.970 ± 1.16 (−4.02, 2.08)	−.521 ± .774 (−2.56, 1.52)	.74
NTS ^{131}I dose > 5.3 mGy?	1567 / 3189	−.824 ± .647 (−2.37, .724)	−.750 ± .656 (−2.44, .941)	.063 ± .681 (−1.73, 1.86)	−2.93 ± 1.54 (−7.00, 1.15)	.053
History of any cancer other than thyroid cancer?	248 / 3186	−.823 ± .646 (−2.37, .724)	−.828 ± .647 (−2.50, .839)	−.686 ± .670 (−2.45, 1.08)	−2.52 ± 2.84 (−10.0, 4.98)	.48
Expanded In-Person Interview?	1212 / 3191	−.824 ± .647 (−2.37, .724)	−.549 ± .640 (−2.20, 1.10)	−.366 ± .796 (−2.47, 1.73)	−.870 ± 1.13 (−3.85, 2.11)	.71

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

Tables IX.L-9 and IX.L-10 display similar results from analyses including history of medical or dental x-ray exposure or of occupational exposure as potential confounding or effect modifying factors. The estimates of the regression coefficient calculated with adjustment for confounding are all close to the unadjusted estimates. Thus there was no evidence that a confounding effect of any of these covariates has obscured a positive dose-response for hyperthyroidism.

There is no evidence of any statistically significant effect modification by any of the covariates in Tables IX.L-9 and IX.L-10, with the possible exception of history of diagnostic x-rays of chest or upper body, including mammograms ($p = 0.033$; Table IX.L-9). The estimated dose-response coefficient was markedly negative (−13.3) for the 352 participants without such histories, but closer to zero for the majority of participants (−0.568). The statistical significance of this difference must be interpreted with caution due to the large number of such comparisons that were performed. Moreover the difference consists of a very negative dose-response in a minority of participants, compared to a less negative coefficient in the remaining participants. Therefore it does not appear that any of the covariates in Tables IX.L-9 and IX.L-10 identified a group in which a clearly significant dose-response was present.

Table IX.L-9. Confounding and Effect Modification by History of Medical and Dental Radiation Exposures: Multinodular Thyroid Gland

Have You Ever Had: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy) Unadjusted	Adjusted for Confounding	Including Effect Modification Group 0	Including Effect Modification Group 1	P
CAT scan of the upper body?	775 / 3149	−.895 ± .659 (−2.47, .684)	−.866 ± .656 (−2.56, .823)	−1.19 ± .791 (−3.28, .895)	.182 ± 1.27 (−3.17, 3.54)	.37
Diagnostic x-rays of the head?	1191 / 3155	−.754 ± .641 (−2.29, .780)	−.704 ± .636 (−2.34, .934)	−.647 ± .835 (−2.85, 1.56)	−.779 ± .98 (−3.36, 1.80)	.62
Diagnostic x-rays of the neck?	966 / 3167	−.813 ± .645 (−2.36, .731)	−.794 ± .643 (−2.45, .864)	−.688 ± .785 (−2.76, 1.38)	−1.00 ± 1.15 (−4.03, 2.03)	.82
Diagnostic x-rays of chest or upper body, including mammograms?	2821 / 3173	−.824 ± .646 (−2.37, .723)	−.785 ± .645 (−2.45, .876)	−13.3 ± 8.01 (−34.4, 7.83)	−.568 ± .62 (−2.21, 1.08)	.033
Diagnostic x-rays of the stomach or mid-back?	692 / 3120	−.885 ± .663 (−2.47, .702)	−.900 ± .662 (−2.61, .807)	−.729 ± .686 (−2.54, 1.08)	−2.28 ± 2.23 (−8.17, 3.60)	.48
Barium enema?	825 / 3159	−.790 ± .644 (−2.33, .752)	−.789 ± .644 (−2.45, .871)	−.473 ± .704 (−2.33, 1.38)	−1.79 ± 1.44 (−5.58, 2.00)	.39
Upper GI?	1146 / 3177	−.818 ± .646 (−2.36, .727)	−.821 ± .645 (−2.48, .841)	−.316 ± .706 (−2.18, 1.55)	−1.98 ± 1.29 (−5.38, 1.41)	.24
Intravenous pyelogram?	398 / 3157	−.896 ± .662 (−2.48, .689)	−.888 ± .663 (−2.59, .819)	−587 ± .658 (−2.32, 1.15)	−4.13 ± 2.86 (−11.7, 3.42)	.16
Fluoroscopy of the upper body?	246 / 3161	−.876 ± .659 (−2.45, .701)	−.880 ± .659 (−2.58, .818)	−.672 ± .657 (−2.41, 1.06)	−4.68 ± 3.96 (−15.1, 5.75)	.22
Nuclear scan (excluding thyroid scan)?	217 / 3162	−.888 ± .660 (−2.47, .692)	−.868 ± .660 (−2.57, .832)	−.727 ± .654 (−2.45, .999)	−6.45 ± 6.18 (−22.8, 9.84)	.24
Dental x-rays that did not usually include a lead shield over the neck area?	1648 / 3191	−.824 ± .647 (−2.37, .724)	−.828 ± .649 (−2.50, .842)	−2.36 ± 1.16 (−5.43, .709)	.134 ± .67 (−1.64, 1.91)	.057

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

Table IX.L-10. Confounding and Effect Modification by Occupational History: Multinodular Thyroid Gland

Have You Ever Worked in Any of the Following? (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted	Adjusted for Confounding	Including Effect Modification		P
				Group 0	Group 1	
Any metal industry?	238 / 3191	–.824 ± .647 (–2.37, .724)	–.814 ± .647 (–2.48, .852)	–.814 ± .647 (–2.52, .892)	--*	--
Any nuclear facility?	371 / 3191	–.824 ± .647 (–2.37, .724)	–.774 ± .649 (–2.45, .897)	–.753 ± .686 (–2.56, 1.06)	–.947 ± 1.98 (–6.17, 4.28)	.93
Any other industry or occupation where you may have been exposed to radioactive materials or x-rays?	442 / 3191	–.824 ± .647 (–2.37, .724)	–.876 ± .651 (–2.55, .802)	–.753 ± .664 (–2.50, .998)	–2.82 ± 3.22 (–11.3, 5.68)	.48
Any of the above industries or occupations?	892 / 3191	–.824 ± .647 (–2.37, .724)	–.746 ± .645 (–2.41, .914)	–.570 ± .674 (–2.35, 1.21)	–1.95 ± 2.02 (–7.29, 3.39)	.49

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.
* Dose-response coefficient not estimable as none of the living evaluable in-area participants with a diagnosis of multinodular thyroid gland ever worked in any metal industry.

Table IX.L-11 displays the results of analyses of possible confounding or effect modification by smoking variables. There was no evidence that the dose-response was significantly confounded by either smoking variable, or that there was a dose-response that differed significantly according to smoking history.

Table IX.L-11. Confounding and Effect Modification by Smoking: Multinodular Thyroid Gland

Have You Ever Smoked Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted	Adjusted for Confounding	Including Effect Modification		P
				Group 0	Group 1	
Cigarettes (unfiltered or filtered)?	1854 / 3183	–.805 ± .647 (–2.35, .744)	–.804 ± .647 (–2.47, .863)	–.053 ± .822 (–2.22, 2.12)	–1.66 ± 1.04 (–4.41, 1.09)	.22
Any of cigarettes, cigar or pipe?	1900 / 3183	–.805 ± .647 (–2.35, .744)	–.804 ± .647 (–2.47, .863)	.036 ± .814 (–2.11, 2.18)	–1.75 ± 1.05 (–4.51, 1.01)	.17

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

L.2.j. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for multinodular thyroid gland are shown in Figure IX.L-1 for each of the 100 dose realizations produced by the CIDER model. The

confidence intervals in that figure are calculated at the 98.33% confidence level, i.e., are adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates. The point estimate of the slope was greater than 0 for only 8 of the 100 realizations, and the confidence interval included 0 for all 100 realizations. Also shown in Figure IX.L-1 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean and mean of each participant's 100 dose realizations. In summary, for none of the 100 realizations of the estimated doses was there a statistically significant dose-response, and for most of the realizations the estimated slope was less than 0.

Figure IX.L-1. Plot of Estimated Slope and 95% CI by Dose Realization: Multinodular Thyroid Gland

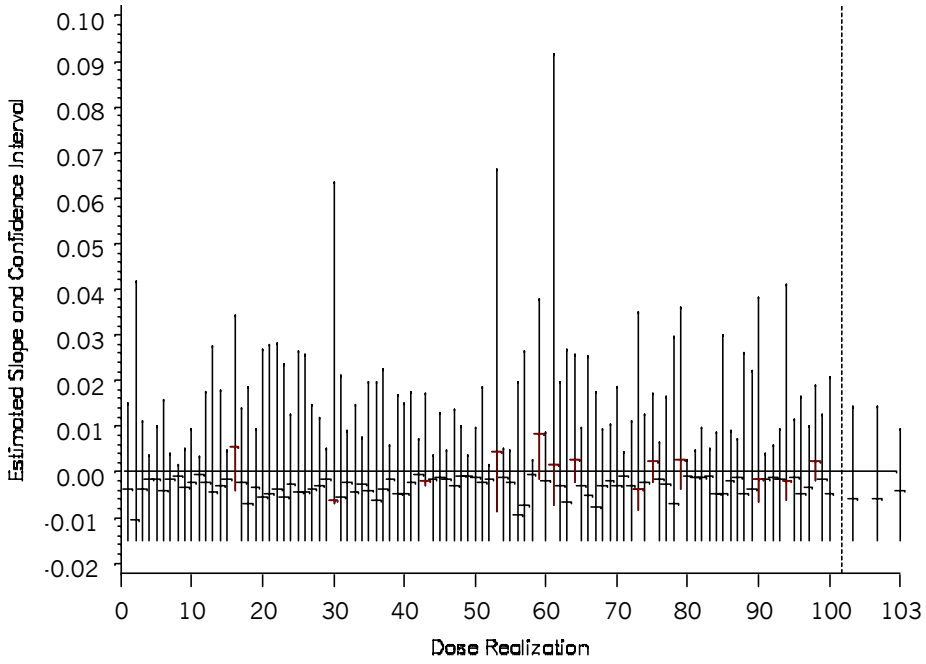
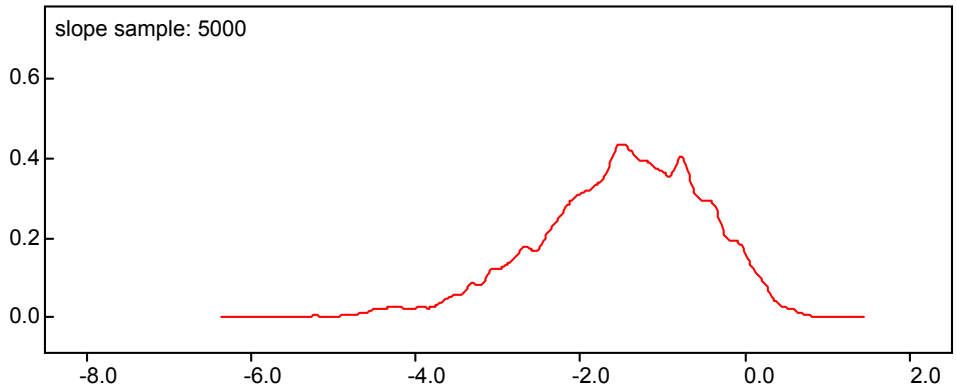


Figure IX.L-2 displays the distribution of the 5000 estimates of the logistic regression coefficient obtained by the simulation procedures described in section VIII.C.2.b.3 above. It is evident from the figure that most of the estimates were between about –5.0 and 1.0. The estimate was less than or equal to 0 for 4770 of the 5000 replications, implying an empirical one-tailed p-value of 0.89. The median estimate was –1.41, and the upper and lower percentiles corresponding to the Bonferroni-adjusted 95% confidence interval were –4.37 and 0.39. These may be compared to the estimate of –0.82 with confidence interval (–2.37, 0.72) obtained using the median dose estimates without adjustment for uncertainty. Thus, this method of adjusting the estimated logistic regression coefficient for the uncertainty in the dose estimates did not provide evidence that the cumulative incidence of multinodular thyroid gland increased with increasing dose.

Figure IX.L-2. Distribution of Simulation Estimates of Logistic Regression Coefficient: Multinodular Thyroid Gland



M. Simple Goiter

M.1. Occurrence of Simple Goiter

The primary and alternative definitions for simple goiter were as follows:

- Primary definition: HTDS evaluation (14 cases)
- Alternative definition #1: HTDS evaluation or medical records (42 cases)
- Alternative definition #2: Any diagnosis or participant/respondent report (70 cases).

The diagnosis of simple goiter was uncommon, with only 14 (0.4%) living evaluable participants having this diagnosis based on HTDS evaluation, 28 (0.8%) based on medical records, and 28 (0.8%) based on a report by the participant or his/her CATI respondent (Table IX.M-1). It should be noted that since this outcome is based solely on physical examination, diagnoses based on medical records are subject to wide variability since exams were done by many different types of providers with differing levels of consistency and frequency and differing criteria for simple goiter. Simple goiter was more commonly diagnosed among women than men.

Table IX.M-1. Basis for Diagnosis of Simple Goiter, by Sex

	Female		Male		Total	
Basis for Diagnosis	No.	%	No.	%	No.	%
Yes	62	3.5	8	0.5	70	2.0
▪ HTDS evaluation	9	0.5	5	0.3	14	0.4
▪ Medical record	27	1.5	1	0.1	28	0.8
▪ Participant/respondent report	26	1.5	2	0.1	28	0.8
No	1684	96.4	1685	99.5	3369	97.9
Unknown	1	0.1	0	--	1	0.0
Total	1747	100.0	1693	100.0	3440	100.0

One living evaluable participant was classified as "unknown" with regard to diagnosis of simple goiter. This participant did not have a medical record indicating simple goiter, but had a participant report of an unknown diagnosis which might have been goiter. This participant was included as a non-case in analyses of the dose-response for simple goiter.

In 30.0% of the cases, simple goiter had one or more of the following etiologies: Graves disease, Hashimoto's thyroiditis, hypothyroidism and/or hyperthyroidism (Table IX.M-2).

Table IX.M-2. Etiologies of Simple Goiter, by Sex

	Female		Male		Total	
Etiology	No.	%	No.	%	No.	%
Graves disease	9	14.5	0	--	9	12.9
Hashimoto's thyroiditis	6	9.7	2	25.0	8	11.4
Hypothyroidism	6	9.7	2	25.0	8	11.4
Hyperthyroidism	1	1.6	0	--	1	1.4
Other	43	69.4	6	75.0	49	70.0
Total with simple goiter	62	100.0	8	100.0	70	100.0

Note: A participant can have >1 etiology

Of those with an other etiology of simple goiter, 44 (89.8%) had no certain etiology, while 4 (8.2%) were due to probable/possible Hashimoto's, and 1 (2.0%) to possible Graves disease (Table IX.M-3).

Table IX.M-3. "Other" Etiologies of Simple Goiter, by Sex

	Female		Male		Total	
Etiology	No.	%	No.	%	No.	%
Uncertain/unknown	39	90.7	5	83.3	44	89.8
Probable/possible Hashimoto's	3	7.0	1	16.7	4	8.2
Probable Graves disease	1	2.3	0	--	1	2.0
Total with an other etiology of simple goiter	43	100.0	6	100.0	49	100.0

M.2. Analysis of Simple Goiter Risk

M.2.a. Primary Analysis

All of the 14 living evaluable participants with a diagnosis of simple goiter based on the HTDS examination, were in-area participants. The proportions with simple goiter are shown by sex, dose category and basis for diagnosis in Table IX.M-4.

Table IX.M-4. Diagnoses of Simple Goiter by Sex, Estimated Dose, and Basis for Diagnosis

A. Female

Thyroid Radiation Dose (mGy)	Living Evaluable Female	Primary Definition: Cases Based on HTDS examination		1st Alternative Definition: Cases Based on HTDS Examination or Medical Records		2nd Alternative Definition: Cases Based on Any Diagnosis or Participant or CATI Report	
	No.	No.	%	No.	%	No.	%
Out of Area	125	0	--	2	1.6	5	4.0
< 10	182	1	0.5	3	1.6	7	3.8
10-49	320	1	0.3	8	2.5	10	3.1
50-99	313	2	0.6	5	1.6	10	3.2
100-149	220	3	1.4	5	2.3	8	3.6
150-199	126	0	--	2	1.6	3	2.4
200-299	139	1	0.7	2	1.4	5	3.6
300-399	144	1	0.7	4	2.8	7	4.9
400-999	171	0	--	5	2.9	7	4.1
1000+	7	0	--	0	--	0	--
Total	1747	9	0.5	36	2.1	62	3.5

B. Male

Thyroid Radiation Dose (mGy	Living Evaluable Male	Primary Definition: Cases Based on HTDS Examination		1st Alternative Definition: Cases Based on HTDS Examination or Medical Records		2nd Alternative Definition: Cases Based on Any Diagnosis or Participant or CATI Report	
	No.	No.	%	No.	%	No.	%
Out of Area	124	0	--	0	--	0	--
< 10	186	1	0.5	1	0.5	2	1.1
10-49	314	1	0.3	1	0.3	1	0.3
50-99	310	1	0.3	2	0.6	3	1.0
100-149	171	1	0.6	1	0.6	1	0.6
150-199	109	0	--	0	--	0	--
200-299	148	0	--	0	--	0	--
300-399	160	0	--	0	--	0	--
400-999	154	1	0.6	1	0.6	1	0.6
1000+	17	0	--	0	--	0	--
Total	1693	5	0.3	6	0.4	8	0.5

Parameter estimates for the linear dose-response model based on the 3191 in-area participants are shown in Table IX.M-5 below. Based on maximum likelihood analysis of the sex-stratified linear probability model, and using the primary dose estimates, the estimated slope B was –0.001 per Gy (row 1 of Table IX.M-5). The lower limit of the Bonferroni-adjusted 95% confidence interval was not estimated, but the upper limit was 0.012 per Gy, and the cumulative incidence of simple goiter did not increase significantly with increasing dose ($p = 0.74$). The corresponding estimated background rates for diagnosis of simple goiter were 0.006 with confidence interval (0.001, 0.011) for women and 0.003 with confidence interval (0, 0.008) for men. When the model was fit by the method of least squares, the estimated slope using either ungrouped or grouped data was even more negative than the maximum likelihood estimate (Table IX.M-5, rows 2 and 3), thereby providing no evidence that risk of simple goiter increased with increasing dose ($p = 0.79$ and 0.70) for the ungrouped and grouped data, respectively.

Table IX.M-5. Summary of Dose-Response Results for Diagnoses of Simple Goiter

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions/ Additional Inclusions	Method of Analysis	Estimates Background Rates Female	Estimates Background Rates Male	Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
1.	Primary definition (HTDS evaluation)	Linear	Primary	None	MLE	.006 ± .002 (.001, .011)	.003 ± .002 (0*, .008)	–.001 ± .008 (NE, .012)	0.74
2.	Primary definition	Linear	Primary	None	LSU	.006 ± .002 (.002, .011)	.004 ± .002 (0*, .009)	–.004 ± .005 (–.017, .008)	0.79
3.	Primary definition	Linear	Primary	None	LSG	.006 ± .002 (.002, .011)	.004 ± .002 (0*, .009)	-.003 ± .006 (-.018, .011)	0.70
4.	Alternative def. #1 (HTDS or medical record)	Linear	Primary	None	MLE	.021 ± .004 (.012, .031)	.004 ± .002 (-.001, .009)	-.002 ± .009 (NE, .019)	0.68
5.	Alternative def. #2 (Any diagnosis or participant/respondent report)	Linear	Primary	None	MLE	.036 ± .005 (.023, .048)	.005 ± .002 (0*, .011)	-.002 ± .011 (NE, .018)	0.74

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model. "0*" indicates that the lower confidence limit for a background rate was less than 0.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.M-5. Summary of Dose-Response Results for Diagnoses of Simple Goiter (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
6.	Primary definition	LQ	Primary	None	LSU	.006 ± .002 (.001, .011)	.004 ± .002 (0*, .009)	Lin: –.005 ± .009 (–.028, .017) Quad: .001 ± .006 (–.014, .016)	Quad: 0.88
7.	Primary definition	Logistic	Primary	None	MLE	.007 (.003, .018)	.004 (.001, .013)	–1.56 ± 1.83 (-5.94, 2.81)	0.83
8.	Primary definition	Linear	Primary	Exclude dose > 1000 mGy	MLE	.006 ± .002 (.001, .011)	.004 ± .002 (0*, .009)	-.004 ± .008 (NE, .014)	0.80
9.	Primary definition	Linear	Primary	Exclude dose > 400 mGy	MLE	.007 ± .003 (0*, .015)	.003 ± .002 (0*, .008)	-.009 ± .016 (NE, .024)	0.84
10.	Primary definition	Linear	Primary	Exclude Ok and F/S geostrata	MLE	.006 ± .002 (.000, .011)	.003 ± .002 (0*, .008)	-.001 ± .008 (NE, .015)	0.70
11.	Primary definition	Linear	Alt. #1	None	MLE	.006 ± .002 (.001, .010)	.003 ± .002 (-.001, .008)	-.001 ± .005 (<-.001, .014)	0.55

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model. "0*" indicates that the lower confidence limit for a background rate was less than 0.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.M-5. Summary of Dose-Response Results for Diagnoses of Simple Goiter (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions/ Additional Inclusions	Method of Analysis	Estimates Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
12.	Primary definition	Linear	Alt. #2	None	MLE	.006 ± .002 (.001, .010)	.003 ± .002 (0*, .008)	-.001 ± .005 (<-.002, .016)	0.56
13.	Primary definition	Linear	Primary	Include OOA (scoping analysis #1)	MLE	.005 ± .002 (.001, .010)	.003 ± .002 (0*, .007)	-.001 ± .008 (NE, .014)	0.71
14.	Primary definition	Linear	Primary	Include OOA (scoping analysis #2)	MLE	.005 ± .002 (.001, .010)	.003 ± .002 (-.001, .007)	-.001 ± .008 (NE, >.014)	0.71

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model. "0*" indicates that the lower confidence limit for a background rate was less than 0.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

M.2.b. Alternative Definitions for Diagnosis of Simple Goiter

Two alternative definitions for cases of simple goiter were considered. The first alternative added 28 cases with diagnoses based on medical records, for a total of 42 cases. The second added another 28 cases based solely on a report from the participant or his/her CATI respondent, for a total of 70 cases. As shown in rows 4 and 5 of Table IX.M-5, for neither alternative definition was there evidence that the cumulative incidence of simple goiter increased with increasing dose ($p = 0.68$ and 0.74 for the first and second alternative analyses, respectively).

M.2.c. Alternative Dose-Response Functions

As shown in row 6 of Table IX.M-5, the estimated regression coefficient for the dose-squared term in the linear-quadratic dose-response model [5] was 0.001 with Bonferroni-adjusted 95% confidence interval ranging from –0.014 to 0.016. Thus the addition of a quadratic term did not significantly improve the fit of the model ($p = 0.88$).

In the analysis of simple goiter based on the HTDS examination, i.e., the primary criterion for defining cases with simple goiter, the regression parameter for the effect of dose in the sex-stratified logistic regression model [2] was estimated as –1.56 with Bonferroni-adjusted 95% confidence interval ranging from –5.94 to 2.81. Thus the cumulative incidence of simple goiter did not increase significantly with increasing dose ($p = 0.83$; Table IX.M-5, row 7).

M.2.d. Effect of Excluding Participants in High Dose Categories

Rows 8 and 9 of Table IX.M-5 show the effect of excluding patients in high dose categories from the analysis of the sex-stratified linear dose-response model. When participants with estimated dose > 1000 mGy were excluded from the analysis, the estimated slope of the dose-response decreased to –0.004 per Gy, providing no evidence that the cumulative incidence of simple goiter increased with increasing dose ($p = 0.80$). Similarly, when participants with estimated dose > 400 mGy were excluded, the estimated slope B was even more negative (–0.009 per Gy), providing no evidence that the cumulative incidence increased with increasing dose ($p = 0.84$).

M.2.e. Effect of Excluding Okanogan and Ferry/Stevens Geostrata

When participants in the Okanogan and Ferry/Stevens geostrata were excluded the estimated slope B was essentially unchanged at –0.001 per Gy, providing no evidence that the cumulative incidence of simple goiter increased significantly with increasing dose ($p = 0.70$; Table IX.M-5, row 10).

M.2.f. Analysis of Simple Goiter in Relation to Alternative Dose Estimates

As shown in rows 11 and 12 of Table IX.M-5, substituting either of the alternative sets of dose estimates for the primary doses caused very little change in the estimated slope of the dose-response. In particular there was no evidence that cumulative incidence of simple goiter increased with increasing doses from either of the alternative dose sets ($p = 0.55$ and $p = 0.56$ for alternative dose sets 1 and 2, respectively).

M.2.g. Scoping Analysis Regarding Out-of-Area Participants

See section VIII.C.1.a.3 for a description of the scoping analyses that were performed to assess the possible impact of including the 249 out-of-area participants. The results of both scoping analyses (Table IX.M-5, rows 13 and 14) were very similar to those of the primary analysis of the in-area participants (Table IX.M-5, row 1). In particular both estimates of the slope were slightly less than zero, providing no evidence that the cumulative incidence of simple goiter increased with increasing dose ($p = 0.71$ for both scoping analyses).

M.2.h. Analysis of Simple Goiter in Relation to Alternative Representations of Exposure

In analyses by geostratum and by dichotomous exposure variable, the sex and age-adjusted comparisons of cumulative incidence were performed as described in section VIII.C.2.a.2.

M.2.h.1. Analysis by Geostratum

There were too few participants (14) with diagnoses of simple goiter (based on the HTDS examination) for a definitive conclusion regarding heterogeneity among the geostrata (see Table IX.M-6). Therefore, the analysis was based on the second alternative definition of simple goiter, i.e., including all diagnoses or reports of simple goiter. The results are shown in Table IX.M-7 below. There was no significant heterogeneity among the nine geostrata ($p=0.26$). The percentages with simple goiter were somewhat higher in the Okanogan and Ferry/Stevens geostrata (5.6% for women, 1.5% for men) than in the remaining geostrata (3.4% and 0.4%), but this heterogeneity between combined geostrata was also not statistically significant ($p=0.095$).

Table IX.M-6. Diagnoses of Simple Goiter Based on the HTDS Evaluation, by Geostratum and Sex

	Female			Male			Total		
Geostratum	No.	Cases	%	No.	Cases	%	No.	Cases	%
Richland	179	1	0.6	173	2	1.2	352	3	0.9
Pasco/Kennewick	508	2	0.4	501	1	0.2	1009	3	0.3
Benton County	376	3	0.8	358	1	0.3	734	4	0.5
Franklin County	73	0	--	76	0	--	149	0	--
Adams County	165	0	--	156	0	--	321	0	--
Walla Walla (city)	133	0	--	131	0	--	264	0	--
Walla Walla County	170	2	1.2	164	0	--	334	2	0.6
Okanogan County	75	0	--	64	1	1.6	139	1	0.7
Ferry/Stevens Counties	68	1	1.5	70	0	--	138	1	0.7
Total	1747	9	0.5	1693	5	0.3	3440	14	0.4

Table IX.M-7. Diagnoses of Simple Goiter Based on Any Source, by Geostratum and Sex

Geostratum	Female No.	Female Cases	Female %	Male No.	Male Cases	Male %	Total No.	Total Cases	Total %
Richland	179	6	3.4	173	2	1.2	352	8	2.3
Pasco/Kennewick	508	22	4.3	501	2	0.4	1009	24	2.4
Benton County	376	13	3.5	358	2	0.6	734	15	2.0
Franklin County	73	3	4.1	76	0	--	149	3	2.0
Adams County	165	2	1.2	156	0	--	321	2	0.6
Walla Walla (city)	133	4	3.0	131	0	--	264	4	1.5
Walla Walla County	170	4	2.4	164	0	--	334	4	1.2
Okanogan County	75	3	4.0	64	1	1.6	139	4	2.9
Ferry/Stevens Counties	68	5	7.4	70	1	1.4	138	6	4.3
Total	1747	62	3.5	1693	8	0.5	3440	70	2.0

M.2.h.2. Analysis by Dichotomous Exposure Variable

See section VIII.B.3.b.2 above for a description of the high and low exposure categories. Only five (0.4%) of the 1257 participants included in these analyses had a diagnosis of simple goiter based on the HTDS examination (see Table IX.M-8). This was too few for a meaningful comparison between the high and low exposure groups. Therefore the analysis was based on the second alternative definition for diagnoses of simple goiter (section IX.M.1, above), i.e., any diagnosis based on HTDS, medical records, or participant or CATI respondent report. As shown in Table IX.M-9, 24 (1.9%) of the 1257 participants had diagnoses of simple goiter based on this alternative criterion, including 9/580 (1.6%) and 15/677 (2.2%) in the high and low exposure groups, respectively. The cumulative incidence of simple goiter based on this alternative definition was not significantly elevated in the high exposure group ($p = 0.75$).

Table IX.M-8. Diagnoses of Simple Goiter Based on HTDS Examination, by Exposure Group and Sex

Exposure Group	Female No.	Female Cases	Female %	Male No.	Male Cases	Male %	Total No.	Total Cases	Total %
Low	351	1	0.3	326	2	0.6	677	3	0.4
High	298	1	0.3	282	1	0.4	580	2	0.3
Total	649	2	0.3	608	3	0.5	1257	5	0.4

Table IX.M-9. Diagnoses of Simple Goiter Based on Any Source, by Exposure Group and Sex

Exposure Group	Female No.	Female Cases	Female %	Male No.	Male Cases	Male %	Total No.	Total Cases	Total %
Low	351	13	3.7	326	2	0.6	677	15	2.2
High	298	8	2.7	282	1	0.4	580	9	1.6
Total	649	21	3.2	608	3	0.5	1257	24	1.9

M.2.i. Confounding and Effect Modification

There were too few participants with diagnoses of simple goiter to permit meaningful analysis of confounding or effect modification.

M.2.j. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for simple goiter are shown in Figure IX.M-1 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in that figure are calculated at the 98.33% confidence level, i.e., are adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates. The point estimate of the slope was greater than zero for only 13 of the 100 realizations, and the confidence interval included zero for all 100 realizations. Also shown in Figure IX.M-1 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean and mean of each participant's 100 dose realizations. In summary, for none of the 100 realizations of the estimated doses was there a statistically significant dose-response, and for most of the realizations the estimated slope was less than 0.

Figure IX.M-1. Plot of Estimated Slope and 95% CI by Dose Realization: Simple Goiter

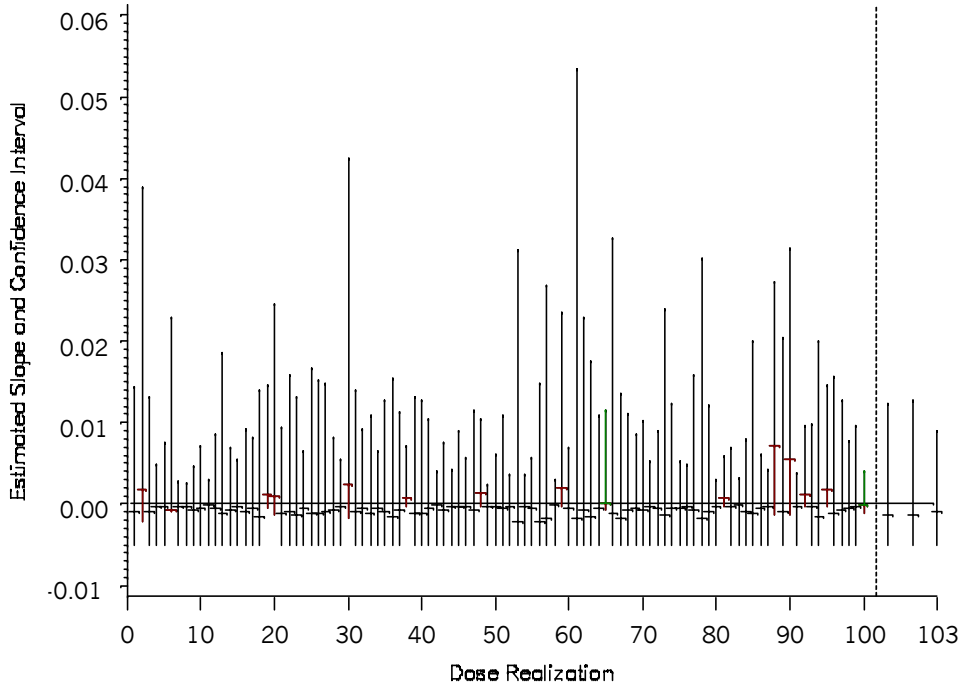
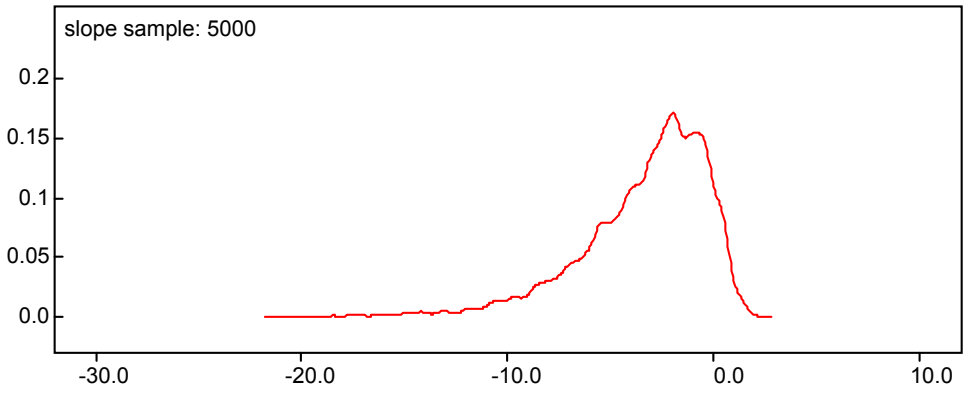


Figure IX.M-2 displays the distribution of the 5000 logistic regression coefficient estimates obtained by the simulation procedures described in section VIII.C.2.b.3 above. It is evident from the figure that most of the estimates were between about –15.0 and 2.0. The estimate was less than or equal to 0 for 4536 of the 5000 replications, implying an empirical one-tailed p-value of 0.91. The median estimate was –2.63, and the upper and lower percentiles corresponding to the Bonferroni-adjusted 95% confidence interval ranging from –14.8 to 1.16. These may be compared to the estimate of –1.56 with confidence interval (–5.94, 2.81) obtained using the median dose estimates without adjustment for uncertainty. Thus, this method of adjusting the estimated logistic regression coefficient for the uncertainty in the dose estimates did not provide evidence that the cumulative incidence of simple goiter increased with increasing dose.

Figure IX.M-2. Distribution of Simulation Estimates of Logistic Regression Coefficient: Simple Goiter



N. Other Thyroid Disease

N.1. Occurrence of Other Thyroid Disease

The primary and alternative definitions for other thyroid disease were as follows:

- Primary definition: HTDS examination or medical records with supporting documentation (4 cases)
- Alternative definition #1: HTDS examination or medical records with or without supporting documentation (6 cases)
- Alternative definition #2: Any diagnosis or participant/respondent report (26 cases).

Four living evaluable participants, all in the in-area group, had diagnoses of other thyroid disease based on their HTDS examinations or medical records with supporting documentation. These included two cases of subacute thyroiditis in women with estimated doses of 342 and 336 mGy; one case of familial thyroglobulin binding deficiency in a male with an estimated dose 102 mGy; and a case of secondary hypothyroidism in a female with an estimated dose 109 mGy.

The first alternative definition added only two cases with diagnoses based on medical records without supporting documentation. Both were cases of subacute thyroiditis in women with estimated doses of 70 and 50 mGy.

For both the primary and first alternative definition of other thyroid disease, there were too few cases for meaningful estimation of the radiation dose-response.

The second alternative definition added 20 participants, primarily with participant or CATI respondent reports of past thyroid disease of unknown type. This brought the total number of cases to 26, of whom four were out-of-area participants. The number of cases and proportions with other thyroid disease are shown by sex and dose category in Table IX.N-1.

Table IX.N-1. Diagnoses of Other Thyroid Disease by Sex, Dose Category, and Basis for Diagnosis

A. Female

Thyroid Radiation Dose	Living Evaluable Female	Primary Definition: Cases Based on HTDS Examination or Medical Records with Supporting Documentation		1st Alternative Definition: Cases Based on HTDS Examination or Medical Records with or without Supporting Documentation		2nd Alternative Definition: Cases Based on Any Diagnosis or Participant or CATI Report	
(mGy)	No.	No.	%	No.	%	No.	%
Out of Area	125	0	--	0	--	1	0.8
< 10	182	0	--	0	--	1	0.5
10-49	320	0	--	0	--	1	0.3
50-99	313	0	--	2	0.6	5	1.6
100-149	220	1	0.5	1	0.5	2	0.9
150-199	126	0	--	0	--	1	0.8
200-299	139	0	--	0	--	2	1.4
300-399	144	2	1.4	2	1.4	2	1.4
400-999	171	0	--	0	--	2	1.2
1000+	7	0	--	0	--	0	--
Total	1747	3	0.2	5	0.3	17	1.0

B. Male

Thyroid Radiation Dose	Living Evaluable Male	Primary Definition: Cases Based on HTDS Examination or Medical Records with Supporting Documentation		1st Alternative Definition: Cases Based on HTDS Examination or Medical Records with or without Supporting Documentation		2nd Alternative Definition: Cases Based on Any Diagnosis or Participant or CATI Report	
(mGy)	No.	No.	%	No.	%	No.	%
Out of Area	124	0	--	0	--	3	2.4
< 10	186	0	--	0	--	1	0.5
10-49	314	0	--	0	--	0	--
50-99	310	0	--	0	--	2	0.6
100-149	171	1	0.6	1	0.6	1	0.6
150-199	109	0	--	0	--	0	--
200-299	148	0	--	0	--	1	0.7
300-399	160	0	--	0	--	0	--
400-999	154	0	--	0	--	1	0.6
1000+	17	0	--	0	--	0	--
Total	1693	1	0.1	1	0.1	9	0.5

Parameter estimates for the linear dose-response model based on the 3191 in-area participants, and using the second alternative definition of other thyroid disease, are shown in Table IX.N-2 below. Based on maximum likelihood analysis of the sex-stratified linear probability model, the estimated slope B was slightly greater than zero (0.002 per Gy) with Bonferroni-adjusted 95% CI ranging from less than –0.002 to 0.024 per Gy, providing no evidence that cumulative incidence increased significantly with increasing dose (one-tailed p = 0.39). The corresponding estimated background rates for diagnosis of other thyroid disease were 0.010 with confidence interval (0.003, 0.016) for women and 0.003 with confidence interval (0, 0.008) for men.

Table IX.N-2. Dose-Response Results for Diagnoses of Other Thyroid Disease

Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
	Female	Male		
Alternative Definition 2: Maximum Likelihood (Any diagnosis or participant/respondent report)	.010 ± .003 (.003, .016)	.003 ± .002 (0*, .008)	.002 ± .007 (< –.002, .024)	0.39

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses. ("<" indicates that the lower confidence limit is less than the indicated value). "0*" indicates that the lower confidence limit for a background rate was less than 0.

In view of the small number of cases, and their heterogeneous and mostly unknown diagnoses, further analyses of this outcome were not performed.

O. Hyperparathyroidism

O.1. Occurrence of Hyperparathyroidism

The primary and alternative definitions of hyperparathyroidism were as follows:

- Primary definition: HTDS evaluation or medical records with supporting documentation (12 cases)
- Alternative definition #1: Any diagnosis or participant/respondent report (14 cases).

Fourteen (0.4%) living evaluable participants had diagnoses of hyperparathyroidism (Table IX.O-1), with 11 being based on the HTDS evaluation, 1 on medical records with supporting documentation, and 2 on a participant or his/her CATI respondent report.

One additional living evaluable participant who did not meet the study's criteria for hyperparathyroidism nevertheless had an elevated calcium in the presence of a high normal PTH level, when the PTH should have been suppressed, highly suggestive of hyperparathyroidism. This participant was included as a case in an additional analysis.

Table IX.O-1. Basis for Diagnosis of Hyperparathyroidism, by Sex

	Female		Male		Total	
Basis for Diagnosis	No.	%	No.	%	No.	%
Yes	10	0.6	4	0.2	14	0.4
▪ HTDS evaluation	9	0.5	2	0.1	11	0.3
▪ Medical records with supporting documentation	1	0.1	0	--	1	0.0
▪ Participant/respondent report	0	--	2	0.1	2	0.1
No	1729	99.0	1687	99.6	3416	99.3
Unknown	8	0.5	2	0.1	10	0.3
Total	1747	100.0	1693	100.0	3440	100.0

Ten living evaluable participants were classified as "unknown" with regard to diagnosis of hyperparathyroidism. These 10 did not have medical record or participant/respondent reports of such diagnoses, and did not have an HTDS evaluation due to the lack of a blood draw (8) or a sufficient amount of blood drawn to determine the serum calcium level (2). These 10 participants were included as non-cases in analyses of the dose-response for hyperparathyroidism.

O.2. Analysis of Hyperparathyroidism Risk

O.2.a. Primary Analysis

Of the 12 living evaluable participants with a diagnosis of hyperparathyroidism based on the HTDS examination or medical records with supporting documentation, one was an out-of-area participant for whom the CIDER program could not calculate a dose estimate. The proportions with hyperparathyroidism are shown by sex, dose category and basis for diagnosis in Table IX.O-2.

Table IX.O-2. Diagnoses of Hyperparathyroidism by Sex, Estimated Dose, and Basis for Diagnosis

A. Female

Thyroid Radiation Dose (mGy)	Living Evaluable Female	Primary Definition: Cases Based on HTDS or Med. Rec. with Supporting Documentation		Alternative Definition: Cases Based on Any Diagnosis or Participant or CATI Report	
	No.	No.	%	No.	%
Out of Area	125	0	--	0	--
< 10	182	1	0.5	1	0.5
10-49	320	2	0.6	2	0.6
50-99	313	2	0.6	2	0.6
100-149	220	1	0.5	1	0.5
150-199	126	1	0.8	1	0.8
200-299	139	2	1.4	2	1.4
300-399	144	1	0.7	1	0.7
400-999	171	0	--	0	--
1000+	7	0	--	0	--
Total	1747	10	0.6	10	0.6

B. Male

Thyroid Radiation Dose (mGy)	Living Evaluable Male	Primary Definition: Cases Based on HTDS or Med. Rec. with Supporting Documentation		Alternative Definition: Cases Based on Any Diagnosis or Participant or CATI Report	
	No.	No.	%	No.	%
Out of Area	124	1	0.8	1	0.8
< 10	186	0	--	0	--
10-49	314	0	--	0	--
50-99	310	0	--	1	0.3
100-149	171	1	0.6	1	0.6
150-199	109	0	--	0	--
200-299	148	0	--	0	--
300-399	160	0	--	0	--
400-999	154	0	--	1	0.6
1000+	17	0	--	0	--
Total	1693	2	0.1	4	0.2

Parameter estimates for the linear dose-response model based on the 3191 in-area participants are shown in Table IX.O-3 below. Based on maximum likelihood analysis of the sex-stratified linear probability model, and using the primary dose estimates, the estimated slope B was –0.000 per Gy (row 1 of Table IX.O-3). The lower limit of the Bonferroni-adjusted 95% confidence interval was not estimated, but the upper limit was 0.013 per Gy, and the cumulative incidence of hyperparathyroidism did not increase significantly with increasing dose ($p = 0.61$). The corresponding estimated background rates for diagnosis of hyperparathyroidism were 0.006 with confidence interval (0, 0.013) for women and 0.001 with confidence interval (0, 0.006) for men. When the model was fit by the method of least squares, the estimated slope using either ungrouped or grouped data was slightly more negative than the maximum likelihood estimate (Table IX.O-3, rows 2 and 3), thereby providing no evidence that risk of hyperparathyroidism increased with increasing dose ($p = 0.74$ and 0.75).

Table IX.O-3. Summary of Dose-Response Results for Diagnoses of Hyperparathyroidism

Row	Outcome	Dose Response Model	Dose Estimates	Exclusion: Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
1.	Primary definition (HTDS evaluation)	Linear	Primary	None	MLE	.006 ± .003 (0*, .013)	.001 ± .002 (0*, .006)	–.000 ± .018 (NE, .013)	0.61
2.	Primary definition	Linear	Primary	None	LSU	.007 ± .002 (.003, .011)	.001 ± .002 (0*, .005)	–.003 ± .005 (–.014, .008)	0.74
3.	Primary definition	Linear	Primary	None	LSG	.007 ± .002 (.003, .011)	.001 ± .002 (0*, .006)	–.004 ± .005 (–.016, .009)	0.75
4.	Alternative def. #1 (Any diagnosis or participant/respondent report)	Linear	Primary	None	MLE	.006 ± .002 (.001, .011)	.002 ± .002 (0*, .006)	.000 ± .006 (< –.001, .021)	0.47
5.	Hyperparathyroidism plus probable case	Linear	Primary	None	MLE	.006 ± .002 (.001, .011)	.003 ± .002 (0*, .008)	–.001 ± .007 (< –.001, .023)	0.54

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model. "0*" indicates that the lower confidence limit for a background rate was less than 0.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.O-3. Summary of Dose-Response Results for Diagnoses of Hyperparathyroidism (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates Female	Estimated Background Rates Male	Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
6.	Primary definition	LQ	Primary	None	LSU	.007 ± .002 (.002, .011)	.001 ± .002 (0*, .006)	Lin: –.003 ± .008 (–.023, .017) Quad: .000 ± .005 (–.013, .013)	Quad: 0.99
7.	Primary definition	Logistic	Primary	None	MLE	.008 (.003, .020)	.001 (.0001, .009)	–1.34 ± 2.00 (-6.14, 3.46)	0.77
8.	Primary definition	Linear	Primary	Exclude dose > 1000 mGy	MLE	.006 ± .003 (0*, .014)	.001 ± .002 (0*, .006)	-.001 ± .018 (NE, .014)	0.67
9.	Primary definition	Linear	Primary	Exclude dose > 400 mGy	MLE	.006 ± .003 (0*, .014)	.000 ± .003 (0*, .006)	.006 ± .022 (<.000, .031)	0.31
10.	Primary definition	Linear	Primary	Exclude Ok and F/S geostrata	MLE	.006 ± .003 (0*, .014)	.001 ± .002 (0*, .006)	-.0003 ± .018 (NE, .013)	0.62
11.	Primary definition	Linear	Alt. #1	None	MLE	.006 ± .003 (0*, .013)	.000 ± .005 (0*, .012)	-.0002 ± .018 (<-.0003, .013)	0.42

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model. "0*" indicates that the lower confidence limit for a background rate was less than 0.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.O-3. Summary of Dose-Response Results for Diagnoses of Hyperparathyroidism

Row	Outcome	Dose Response Model	Dose Estimates	Exclusion: Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
12.	Primary definition	Linear	Alt. #2	None	MLE	.006 ± .003 (0*, .013)	.001 ± .001 (0*, .004)	-.0003 ± .011 (NE, .013)	0.60
13.	Primary definition	Linear	Primary	Include OOA (scoping analysis #1)	MLE	.006 ± .002 (.0002, .011)	.001 ± .001 (0*, .004)	-.0005 ± .011 (NE, .011)	0.68
14.	Primary definition	Linear	Primary	Include OOA (scoping analysis #2)	MLE	.006 ± .002 (.000, .011)	.001 ± .001 (0*, .004)	-.0005 ± .011 (NE, >.010)	0.68

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model. "0*" indicates that the lower confidence limit for a background rate was less than 0.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

O.2.b. Alternative Definitions for Diagnosis of Hyperparathyroidism

One alternative definition for cases of hyperparathyroidism was considered. This alternative added two cases based solely on reports from the participant or his/her CATI respondent, bringing the total to 14 cases. The two added cases had estimated doses of 475 and 92 mGy. As shown in row 4 of Table IX.O-3 above, in the alternative analysis the cumulative incidence of hyperparathyroidism did not increase significantly with increasing dose (p = 0.47), with an estimated slope of 0.000 per Gy and Bonferroni-adjusted 95% confidence interval ranging from less than –0.001 to 0.021 per Gy.

O.2.b.1 Effect of Including Probable Diagnoses of Hyperparathyroidism

As described in IX.O.1 above, one living evaluable participant who wasn't counted as a case of hyperparathyroidism might truly have been a case. This participant was not counted as a case in the primary analysis of hyperparathyroidism to avoid introducing a possible reporting bias. However in view of the importance of hyperparathyroidism as a disease outcome, an additional analysis in which this participant was counted as a case was performed. This participant was in the in-area group, with an estimated thyroid radiation dose of 159 mGy. Counting this participant as a case rather than a noncase in the dose-response analysis had almost no impact on the results (Table IX.O-3, row 5): the estimated slope of the dose-response was slightly less than zero (–0.001 per Gy, with confidence interval ranging from less than –0.001 to 0.023 per Gy) providing no evidence of a positive dose-response (p=0.54).

O.2.c. Alternative Dose-Response Functions

As shown in row 6 of Table IX.O-3, the estimated regression coefficient for the dose-squared term in the linear-quadratic dose-response model [5] was 0.000 with Bonferroni-adjusted 95% confidence interval ranging from –0.013 to 0.013. Thus the addition of a quadratic term did not significantly improve the fit of the model (p = 0.99).

The regression parameter for the effect of dose in the sex-stratified logistic regression model [2] was estimated as –1.34, with Bonferroni-adjusted 95% confidence limits –6.14 and 3.46, indicating that the cumulative evidence of hyperparathyroidism did not increase significantly with increasing dose (p = 0.77; see row 7 of Table IX.O-3).

O.2.d. Effect of Excluding Participants in High Dose Categories

As shown in rows 8 and 9 of Table IX.O-3, when participants in high dose categories were excluded, the cumulative incidence of hyperparathyroidism did not increase significantly with increasing dose (p = 0.67 and p = 0.31 when those with doses >1000 mGy and >400 mGy were excluded, respectively).

O.2.e. Effect of Excluding Okanogan and Ferry/Stevens Geostrata

When participants in the Okanogan and Ferry/Stevens were excluded, the estimated slope B was slightly less than zero (-0.0003 with Bonferroni-adjusted 95% CI upper confidence limit of 0.013), providing no evidence that the cumulative incidence of hyperparathyroidism increased with increasing dose (p = 0.62; Table IX.O-3, row 10).

O.2.f. Analysis of Hyperparathyroidism in Relation to Alternative Dose Estimates

As shown in rows 11 and 12 of Table IX.O-3 above, for neither set of alternative dose estimates did the cumulative incidence of hyperparathyroidism increase significantly with increasing dose ($p = 0.42$ and 0.60 for dose alternatives 1 and 2, respectively).

O.2.g. Scoping Analysis Regarding Out-of-Area Participants

See section VIII.C.1.a.3 for a description of the scoping analyses that were performed to assess the possible impact of including the 249 out-of-area participants. As shown in rows 13 and 14 of Table IX.O-3, the results of both scoping analyses were very similar to those of the primary analysis (row 1). In particular, neither scoping analysis provided evidence that the risk of hyperparathyroidism increased with increased thyroid dose ($p = 0.68$ for both scoping analyses).

O.2.h. Analysis of Hyperparathyroidism in Relation to Alternative Representations of Exposure

In the analyses by geostratum and by dichotomous exposure variable, the sex and age-adjusted comparisons of cumulative incidence were performed as described in section VIII.C.2.a.2.

O.2.h.1. Analysis by Geostratum

Since there were only 12 participants with hyperparathyroidism diagnosed according to the primary definition (HTDS or medical records with documentation), the test for heterogeneity among the 9 geostrata had little statistical power. Therefore the absence of significant heterogeneity ($p = 0.71$) was not strong evidence against the possibility that the cumulative incidence of hyperparathyroidism might in fact vary among the geostrata.

Table IX.O-4. Diagnoses of Hyperparathyroidism Based on the HTDS Evaluation or on Medical Records with Supporting Documentation, by Geostratum and Sex

	Female			Male			Total		
Geostratum	No.	Cases	%	No.	Cases	%	No.	Cases	%
Richland	179	0	--	173	0	--	352	0	--
Pasco/Kennewick	508	4	0.8	501	1	0.2	1009	5	0.5
Benton County	376	3	0.8	358	0	--	734	3	0.4
Franklin County	73	0	--	76	0	--	149	0	--
Adams County	165	0	--	156	1	0.6	321	1	0.3
Walla Walla (city)	133	1	0.8	131	0	--	264	1	0.4
Walla Walla County	170	1	0.6	164	0	--	334	1	0.3
Okanogan County	75	0	--	64	0	--	139	0	--
Ferry/Stevens Counties	68	1	1.5	70	0	--	138	1	0.7
Total	1747	10	0.6	1693	2	0.1	3440	12	0.3

O.2.h.2. Analysis by Dichotomous Exposure Variable

See section VIII.B.3.b.2 above for a description of the high and low exposure categories. Only six (0.5%) of the 1257 participants included in these analyses had a diagnosis of hyperparathyroidism based on the HTDS examination. Therefore the comparison between the high and low exposure groups was based on the alternative definition for diagnoses of hyperparathyroidism (section IX.O.1 above), i.e., any diagnosis based on the HTDS examination, medical records, or participant or CATI respondent report. As shown in Table IX.O-5 below, using the alternative criterion increased the number of cases to only seven (0.6%). Four of these cases occurred among the 580 participants in the high exposure category (0.7%), compared to three (0.4%) among 677 participants in the low exposure group. Consequently the cumulative incidence of hyperparathyroidism was not significantly elevated in the high exposure group (p = 0.43).

Table IX.O-5. Diagnoses of Hyperparathyroidism Based on HTDS Any Diagnosis or Participant/CATI Respondent Report, by Exposure Group and Sex

	Female			Male			Total		
Exposure Group	No.	Cases	%	No.	Cases	%	No.	Cases	%
Low	351	2	0.6	326	1	0.3	677	3	0.4
High	298	2	0.7	282	2	0.7	580	4	0.7
Total	649	4	0.6	608	3	0.5	1257	7	0.6

O.2.i. Confounding and Effect Modification

There were too few participants with diagnoses of hyperparathyroidism to permit meaningful analysis of confounding or effect modification.

O.2.j. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for hyperparathyroidism are shown in Figure IX.O-1 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in that figure are calculated at the 98.33% confidence level, i.e., are adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates. The point estimate of the slope was greater than zero for 51 of the 100 realizations, and the confidence interval included zero for 97 of the 100 realizations. Also shown in Figure IX.O-1 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean and mean of each participant's 100 dose realizations. In summary, for only three of the 100 realizations of the estimated doses was there a statistically significant dose-response, and for about half of the realizations the estimated slope was less than 0.

Figure IX.O-1. Plot of Estimated Slope and 95% CI by Dose Realization: **Hyperparathyroidism**

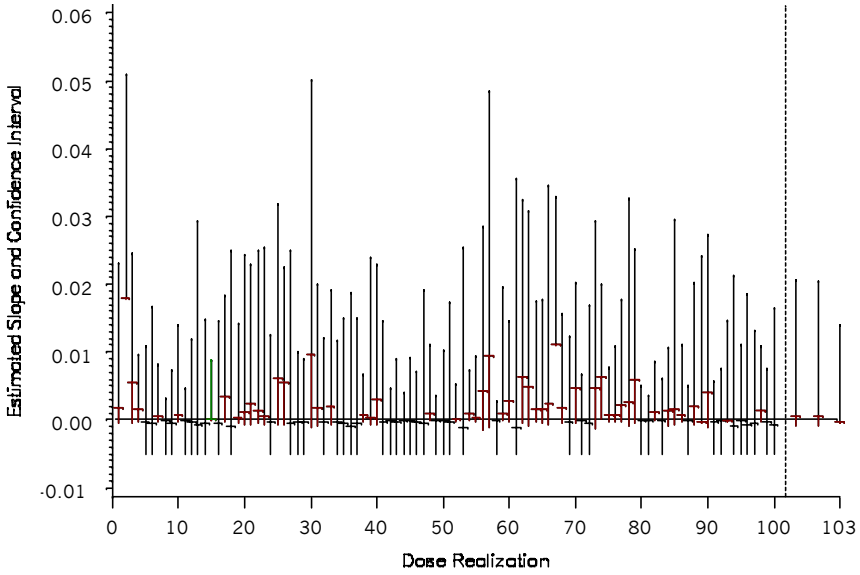
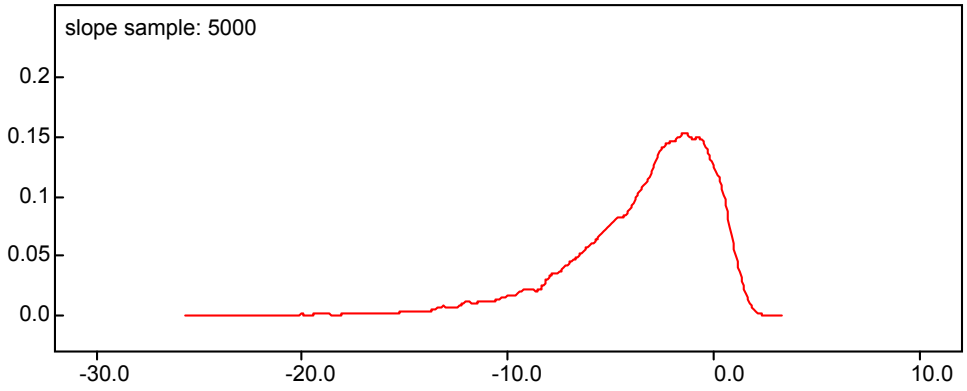


Figure IX.O-2 displays the distribution of the 5000 estimates of the logistic regression coefficient obtained by the simulation procedures described in section VIII.C.2.b.3 above. It is evident from the figure that most of the estimates were between about –15.0 and 2.0. The estimate was less than or equal to 0 for 4378 of the 5000 replications, implying an empirical one-tailed p-value of 0.88. The median estimate was –2.56, and the upper and lower percentiles corresponding to the Bonferroni-adjusted 95% confidence interval ranging from –15.9 to 1.42. These may be compared to the estimate of –1.34 with confidence interval (–6.14, 3.46) obtained using the median dose estimates without adjustment for uncertainty. Thus, this method of adjusting the estimated logistic regression coefficient for the uncertainty in the dose estimates did not provide evidence that the cumulative incidence of hyperparathyroidism increased with increasing dose.

Figure IX.O-2. Distribution of Simulation Estimates of Logistic Regression Coefficient: Hyperparathyroidism



P. Ultrasound-Detected Abnormalities of the Thyroid (Thyroid UDAs)

P.1. Occurrence of Ultrasound-Detected Abnormalities of the Thyroid

The thyroid gland was visible in the ultrasound examinations of 3429 of the 3440 living evaluable participants. For 11 participants the thyroid was not visible, 10 because of thyroid surgery and one because the sonographer couldn't adequately visualize the thyroid. Among the 3429 whose thyroids were visible, 1596 (46.5%) had one or more ultrasound-detected abnormalities (Table IX.P-1), including palpable thyroid UDAs (224 or 6.5%), nonpalpable focal thyroid UDAs (1309 or 38.2%), and diffuse thyroid UDAs (458 or 13.4%). All three types of thyroid UDA were more frequent among women than men. Ultrasound-detected thyroid abnormalities were based only on the HTDS evaluation, not on any prior ultrasounds.

Table IX.P-1. Ultrasound-Detected Abnormalities, by Sex and Type of Abnormality

	Female		Male		Total	
Ultrasound Finding	No.	%	No.	%	No.	%
Thyroid gland visible on ultrasound	1738	100.0	1691	100.0	3429	100.0
Normal ultrasound	774	44.5	1059	62.6	1833	53.5
Any abnormality	964	55.5	632	37.4	1596	46.5
Palpable thyroid UDAs	154	8.9	70	4.1	224	6.5
Nonpalpable focal thyroid UDAs	784	45.1	525	31.0	1309	38.2
Diffuse thyroid UDAs	306	17.6	152	9.0	458	13.4

Note: a participant can have more than one of palpable, nonpalpable focal and diffuse thyroid UDAs.

P.1.a. Additional Outcomes Related to Ultrasound-Detected Abnormalities of the Thyroid

P.1.a.1 Thyroid UDAs by Size

To determine whether the size of thyroid UDAs detected increased in relation to estimated dose, three additional outcomes were defined. These included the presence of a focal thyroid UDA with maximum dimension at least 5 mm, the presence of a focal thyroid UDA with maximum dimension of at least 10 mm, and the presence of a focal thyroid UDA with average dimension of at least 15 mm. Among the 3429 participants whose thyroids were visible, 1142 (33.3%) had a focal thyroid UDA with maximum dimension ≥ 5 mm, 622 (18.1%) had a focal thyroid UDA with maximum dimension ≥ 10 mm, and 166 (4.8%) had a focal thyroid UDA with average dimension ≥ 15 mm (Table IX.P-2).

Table IX.P-2. Ultrasound-Detected Abnormalities, by Sex and Size of Abnormality

	Female		Male		Total	
Ultrasound Findings	No.	%	No.	%	No.	%
Thyroid gland visible on ultrasound	1738	100.0	1691	100.0	3429	100.0
UDA ≥ 5 mm maximum dimension	701	40.3	441	26.1	1142	33.3
UDA ≥ 10 mm maximum dimension	390	22.4	232	13.7	622	18.1
UDA ≥ 15 mm average dimension	105	6.0	61	3.6	166	4.8

P.2. Analysis of Any Ultrasound-Detected Abnormality Risk

P.2.a. Primary Analysis

Among the 3429 participants whose thyroids were visible on ultrasound, 1596 (46.5%) had some type of ultrasound-detected abnormality. These included 3181 in-area participants, of whom 1481 (46.6%)

had any thyroid UDAs, and 248 out of area participants, of whom 115 (46.4%) had thyroid UDAs. The proportions with any thyroid UDA are shown by sex and dose category in Table IX.P-3. The prevalence of thyroid UDAs was higher among women (55.5%) compared to men (37.4%). The numbers and proportions of cases of additional outcomes related to UDAs are also shown in Table IX.P-3.

Table IX.P-3. Any Ultrasound-Detected Abnormality by Sex and Dose Category

A. Female

Thyroid Radiation Dose	L.E. with Ultrasound	Any Thyroid UDA		Focal Thyroid UDA with Maximum Dimension ≥ 5mm		Focal Thyroid UDAs with Maximum Dimension ≥ 10 mm		Focal Thyroid UDAs with Average Dimension ≥ 15 mm	
(mGy)	No.	No.	%	No.	%	No.	%	No.	%
Out of Area	124	64	51.6	50	40.3	26	21.0	4	3.2
< 10	182	100	54.9	73	40.1	43	23.6	10	5.5
10-49	318	171	53.8	126	39.6	73	23.0	16	5.0
50-99	311	172	55.3	118	37.9	65	20.9	16	5.1
100-149	220	131	59.5	97	44.1	53	24.1	13	5.9
150-199	125	65	52.0	49	39.2	32	25.6	14	11.2
200-299	137	79	57.7	59	43.1	26	19.0	6	4.4
300-399	143	80	55.9	55	38.5	26	18.2	11	7.7
400-999	171	100	58.5	73	42.7	46	26.9	15	8.8
1000+	7	2	28.6	1	14.3	0	--	0	--
Total	1738	964	55.5	701	40.3	390	22.4	105	6.0

L.E. = living evaluable participants

B. Male

Thyroid Radiation Dose	L.E. with Ultrasound	Any Thyroid UDA		Focal Thyroid UDA with Maximum Dimension ≥ 5mm		Focal Thyroid UDA with Maximum Dimension ≥ 10 mm		Focal Thyroid UDA with Average Dimension ≥ 15 mm	
(mGy)	No.	No.	%	No.	%	No.	%	No.	%
Out of Area	124	51	41.1	39	31.5	17	13.7	2	1.6
< 10	185	72	38.9	52	28.1	29	15.7	9	4.9
10-49	314	111	35.4	78	24.8	47	15.0	14	4.5
50-99	310	103	33.2	72	23.2	40	12.9	12	3.9
100-149	171	64	37.4	40	23.4	23	13.5	4	2.3
150-199	109	46	42.2	34	31.2	16	14.7	5	4.6
200-299	148	66	44.6	53	35.8	29	19.6	9	6.1
300-399	160	65	40.6	40	25.0	18	11.3	4	2.5
400-999	153	44	28.8	27	17.6	10	6.5	1	0.7
1000+	17	10	58.8	6	35.3	3	17.6	1	5.9
Total	1691	632	37.4	441	26.1	232	13.7	61	3.6

L.E. = living evaluable participants

Parameter estimates for the linear dose-response model based on the 3181 in-area participants with ultrasound results are shown in Table IX.P-4. Based on maximum likelihood analysis of the sex-stratified linear probability model, the risk of having any type of thyroid UDA did not increase significantly with

estimated dose (p = 0.21), with an estimated slope B of 0.031 per Gy, and 95% CI ranging from –0.059 to 0.116 per Gy (Table IX.P-4, row 1). Estimation by least squares using the ungrouped data gave nearly identical results, and the least squares fit to the grouped data were similar (Table IX.P-4, rows 2 and 3).

Table IX.P-4. Dose-Response Results for Diagnoses of Any Thyroid UDA

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
1.	Primary definition	Linear	Primary	None	MLE	.552 ± .014 (.519, .586)	.365 ± .014 (.332, .399)	.031 ± .038 (–.059, .116)	0.21
2.	Primary definition	Linear	Primary	None	LSU	.552 ± .014 (.519, .585)	.365 ± .014 (.331, .399)	.032 ± .039 (–.061, .125)	0.21
3.	Primary definition	Linear	Primary	None	LSG	.556 ± .014 (.522, .591)	.369 ± .015 (.334, .405)	.008 ± .045 (–.099, .115)	0.43
4.	Focal thyroid UDA with max dimension ≥ 5 mm	Linear	Primary	None	MLE	.406 ± .014 (.373, .438)	.259 ± .013 (.228, .290)	-.013 ± .037 (-.097, .077)	0.64
5.	Focal thyroid UDA with max dimension ≥ 10 mm	Linear	Primary	None	MLE	.231 ± .011 (.204, .258)	.143 ± .010 (.119, .167)	-.033 ± .026 (<-.061, .038)	0.88

Entries in the tables are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.P-4. Dose-Response Results for Diagnoses of Any Thyroid UDA (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
6.	Focal thyroid UDA with average dimension ≥ 15 mm	Linear	Primary	None	MLE	.063 ± .007 (.047, .079)	0.038 ± .005 (.025, .051)	-.001 ± .015 (<-.017, .044)	0.53
7.	Primary definition	LQ	Primary	None	LSU	.546 ± .015 (.509, .584)	.359 ± .015 (.321, .397)	Lin: .086 ± .067 (–.078, .250) Quad: –.045 ± .044 (–.153, .064)	Quad: 0.30
8.	Primary definition	Logistic	Primary	None	MLE	.552 (.518, .586)	.365 (.333, .399)	.133 ± .162 (–.254, .520)	0.21
9.	Primary definition	Linear	Primary	Exclude dose > 1000 mGy	MLE	.552 ± .015 (.517, .587)	.362 ± .015 (.327, .397)	.042 ± .049 (<–.075, >.159)	0.20
10.	Primary definition	Linear	Primary	Exclude dose > 400 mGy	MLE	.535 ± .016 (.496, .575)	.356 ± .016 (.317, .395)	.179 ± .086 (–.027, .384)	0.019
11.	Primary definition	Linear	Primary	Exclude OK and F/S geostrata	MLE	.547 ± .015 (.512, .583)	.353 ± .015 (.318, .387)	.047 ± .038 (–.045, .130)	0.11

Entries in the tables are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.P-4. Dose-Response Results for Diagnoses of Any Thyroid UDA (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
12.	Primary definition	Linear	Alt. #1	None	MLE	.556 ± .014 (.521, .590)	.369 ± .014 (.335, .403)	.009 ± .038 (–.080, .097)	0.40
13.	Primary definition	Linear	Alt. #2	None	MLE	.555 ± .014 (.521, .590)	.368 ± .014 (.335, .402)	.012 ± .038 (–.078, .103)	0.37
14.	Primary definition	Linear	Primary	Include OOA (scoping analysis #1)	MLE	.549 ± .013 (.518, .581)	.368 ± .013 (.337, .400)	.033 ± .037 (<-.056, .017)	0.19
15.	Primary definition	Linear	Primary	Include OOA (scoping analysis #2)	MLE	.550 ± .013 (.518, .582)	.369 ± .013 (.338, .401)	.027 ± .037 (–.062, .112)	0.24

Entries in the tables are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

P.2.b. Effect of Using Alternative Size Criteria for Thyroid UDAs

To assess whether the dose-response results might be affected by the size of focal thyroid UDAs, three additional outcomes were analyzed. These included the presence of a focal thyroid UDA with maximum dimension at least 5 mm, the presence of a focal thyroid UDA with maximum dimension at least 10 mm, and the presence of a focal thyroid UDA with average dimension at least 15 mm. These additional analyses applied only to palpable and nonpalpable focal thyroid UDAs, since diffuse thyroid UDAs were not defined by any size criterion. In none of these additional analyses was there any evidence that the risk of having a focal thyroid UDA of a particular size increased with increasing dose ($p = 0.64$, 0.88 and 0.53 for the presence of focal thyroid UDA with maximum dimension of 5 mm, maximum dimension of 10 mm and average dimension of 15 mm, respectively; Table IX.P-4, rows 4, 5, and 6).

P.2.c. Alternative Dose-Response Functions

As shown in row 7 of Table IX.P-4, the estimated regression coefficient for the dose-squared term in the linear-quadratic dose-response model [5] was –0.045 with Bonferroni-adjusted 95% confidence interval ranging from –0.153 to 0.064. Thus the addition of a quadratic term did not significantly improve the fit of the model ($p = 0.30$).

The regression parameter for the effect of dose in the sex-stratified logistic regression model [2] was estimated as 0.133, with Bonferroni-adjusted 95% confidence limits –0.254 and 0.520, indicating that the prevalence of any thyroid UDA did not increase significantly with increasing dose ($p = 0.21$; Table IX.P-4, row 8).

P.2.d. Effect of Excluding Participants in High Dose Categories

In the analyses excluding participants with estimated dose > 1000 mGy (Table IX.P-4, row 9), the estimated slope B was not significantly greater than zero (0.042 per Gy, with Bonferroni-adjusted 95% CI ranging from less than –0.075 to greater than 0.159) providing no evidence that the prevalence of any thyroid UDA increased with increasing dose ($p = 0.20$). When all participants with estimated dose > 400 mGy were excluded (Table IX.P-4, row 10), the estimated slope B was not significantly greater than zero (0.179 per Gy, with Bonferroni-adjusted 95% CI ranging from –0.027 to 0.384). Although there was some evidence that the prevalence of any thyroid UDA increased with increasing dose ($p = .019$), this finding was not considered statistically significant given the large number of such tests that were performed.

P.2.e. Effect of Excluding Okanogan and Ferry/Stevens Geostrata

As shown in row 11 of Table IX.P-4, excluding Okanogan and Ferry/Stevens slightly reduced the estimated background rates for both men and women. The reductions are small, because the Okanogan and Ferry/Stevens geostrata account for only 255 (8.0%) of the 3181 in-area living evaluable participants with ultrasound results. As a result of these reductions in the background rates and the fact that Okanogan and Ferry/Stevens geostrata tend to have low doses, the estimated slope increased slightly, from 0.031 to 0.047, and the statistical significance of the dose-response changed from $p = 0.21$ to $p = 0.11$.

P.2.f. Analysis of Ultrasound-Detected Abnormalities in Relation to Alternative Dose Estimates

As shown in row 12 of Table IX.P-4, using the first alternative dose estimates, the estimated slope B was not significantly greater than zero (0.009 per Gy with Bonferroni-adjusted 95% CI ranging from

-0.080 to 0.097), providing no evidence that prevalence increased with increasing dose (p = 0.40; Table IX.P-4, row 12). Similar results were found with the second set of alternative dose estimates (Table IX.P-4, row 13).

P.2.g. Scoping Analysis Regarding Out-of-Area Participants

See section VIII.C.1.a.3 for a description of the scoping analyses that were performed to assess the possible impact of the 249 out-of-area participants. The results of both scoping analyses were virtually the same as the primary analysis and provided no evidence that the prevalence of any thyroid UDA increased with increasing dose (Table IX.P-4, rows 14 and 15).

P.2.h. Analysis of Any Thyroid UDAs In Relation to Alternative Representations of Exposure

In the analyses by geostratum and by dichotomous exposure variable, the sex and age-adjusted comparisons of prevalence were performed as described in section VIII.C.2.a.2.

P.2.h.1. Analysis by Geostratum

As shown in Table IX.P-5, among the 3429 living evaluable in area or out-of-area participants with ultrasound results, the proportions with any UDAs ranged from 83/131 (63.4% in the Walla Walla City geostratum) to 92/177 (52.0%, Richland) for women, and from 32/63 (50.8%, Okanogan County) to 41/164 (25.0%, Walla Walla County) for men (p = 0.014 for heterogeneity among the nine geostrata). In particular the percentages with any UDAs were somewhat higher in the Okanogan and Ferry/Stevens geostrata (58.7% for women, 48.1% for men) than in the remaining geostrata (55.2% and 36.5%, respectively; p = 0.012). Since it was likely that participants in the Okanogan and Ferry/Stevens geostrata tended to have lower thyroid doses from Hanford's ^{131}I than those in other geostrata, it does not appear that these differences can be attributed to an effect of Hanford's ^{131}I.

Table IX.P-5. Any Ultrasound-Detected Abnormality, by Geostratum and Sex

	Female			Male			Total		
Geostratum	No.	Cases	%	No.	Cases	%	No.	Cases	%
Richland	177	92	52.0	172	60	34.9	349	152	43.6
Pasco/Kennewick	505	273	54.1	501	176	35.1	1006	449	44.6
Benton County	375	206	54.9	358	146	40.8	733	352	48.0
Franklin County	73	42	57.5	76	36	47.4	149	78	52.3
Adams County	165	93	56.4	156	66	42.3	321	159	49.5
Walla Walla (city)	131	83	63.4	131	43	32.8	262	126	48.1
Walla Walla County	169	91	53.8	164	41	25.0	333	132	39.6
Okanogan County	75	43	57.3	63	32	50.8	138	75	54.3
Ferry/Stevens Counties	68	41	60.3	70	32	45.7	138	73	52.9
Total	1738	964	55.5	1691	632	37.4	3429	1596	46.5

P.2.h.2. Analysis by Dichotomous Exposure Variable

Ultrasound was not evaluable for 2 of the 1257 living evaluable participants included in these analyses. Of the 1255 participants included in these analyses, 611 (48.7%) had one or more thyroid UDAs (Table IX.P-6). These included 291/580 (50.2%) in the high exposure group, and 320/675 (47.4%) in the low exposure group. Based on the logistic regression analysis with adjustment for the effect of sex and age

at HTDS examination, the proportion of participants with any thyroid UDA was not significantly elevated in the high exposure group (p = 0.11).

Table IX.P-6. Any Ultrasound-Detected Abnormality, by Exposure Group and Sex

Exposure	Female			Male			Total		
Group	No.	Cases	%	No.	Cases	%	No.	Cases	%
Low	350	192	54.9	325	128	39.4	675	320	47.4
High	298	183	61.4	282	108	38.3	580	291	50.2
Total	648	375	57.9	607	236	38.9	1255	611	48.7

P.2.i. Confounding and Effect Modification

As described in section VIII above, additional sex-stratified logistic regression models were investigated to examine the possibility that the primary dose-response results for any thyroid UDA might be influenced by confounding, and to search for factors that might modify a radiation dose-response. Table IX.P-7 displays results for models including sex, age at first exposure to Hanford I-131 (prenatal, or < 180 days), age at HTDS examination, estimated dose from the NTS, history of any cancer other than thyroid, and HTDS interview type.

Note that sex was not analyzed as a possible confounder since its effect was already adjusted for in the sex-stratified model. None of the other factors in Table IX.P-7 appears to be a confounder: for none does the adjusted estimate of the regression coefficient differ markedly from the unadjusted estimate. Therefore, it does not appear that omitting these factors introduced any important bias in the dose-response results.

Table IX.P-7. Confounding and Effect Modification by Sex, Age at Exposure or HTDS Examination, Estimated Thyroid Dose from NTS, History of Any Cancer Other than Thyroid, and Interview Type: Any Ultrasound-Detected Abnormality

		Estimated Dose-Response Coefficient (per Gy)				
Covariate (0=No, 1=Yes)	Yes / Total	Unadjusted Estimate	Incl. Confounding Estimate	Including Effect Modification Group 0	Including Effect Modification Group 1	P
Female?	1614 / 3181	.133 ± .162 (–.254, .520)	Not Applicable	.198 ± .226 (–.368, .763)	.067 ± .228 (–.504, .637)	.68
Prenatal exposure?	1031 / 3181	.133 ± .162 (–.254, .520)	.087 ± .163 (–.332, .506)	.190 ± .188 (–.305, .685)	–.235 ± .333 (–1.12, .644)	.26
1st exposure before age 180 days?	1474 / 3181	.133 ± .162 (–.254, .520)	.121 ± .162 (–.297, .539)	.373 ± .279 (–.362, 1.11)	–.010 ± .200 (–.538, .519)	.26
Age at exam > 50?	1993 / 3181	.133 ± .162 (–.254, .520)	.178 ± .164 (–.246, .601)	–.038 ± .297 (–.822, .746)	.271 ± .197 (–.248, .791)	.39
NTS I-131 dose > 5.3 mGy?	1563 / 3179	.127 ± .162 (–.260, .514)	.111 ± .165 (–.314, .536)	.106 ± .219 (–.471, .682)	.118 ± .251 (–.544, .781)	.97
History of any cancer other than thyroid?	248 / 3176	.138 ± .162 (–.249, .525)	.141 ± .162 (–.276, .557)	.219 ± .176 (–.244, .683)	–.300 ± .427 (–1.43, .827)	.25
Expanded In-Person Interview?	1205 / 3181	.133 ± .162 (–.254, .520)	.159 ± .165 (–.266, .584)	.277 ± .263 (–.416, .970)	.083 ± .211 (–.475, .640)	.56

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

Tables IX.P-8 and IX.P-9 display similar results from analyses including history of medical or dental x-ray exposure or occupational exposure as potential confounding or effect modifying factors. Specifically, none of the factors in these tables appears to be a confounder or an effect modifier.

Table IX.P-8. Confounding and Effect Modification by History of Medical and Dental Radiation Exposures: Any Ultrasound-Detected Abnormality

Have You Ever Had: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted Estimate	Incl. Confounding Estimate	Including Effect Modification		
				Group 0	Group 1	P
CAT scan of the upper body?	775 / 3139	.130 ± .162 (–.257, .518)	.130 ± .162 (–.288, .548)	.064 ± .178 (–.406, .533)	.450 ± .391 (–.581, 1.48)	.37
Diagnostic x-rays of the head?	1188 / 3145	.133 ± .163 (–.257, .523)	.129 ± .163 (–.291, .549)	.257 ± .204 (–.280, .795)	–.110 ± .279 (–.845, .625)	.29
Diagnostic x-rays of the neck?	960 / 3157	.126 ± .162 (–.263, .515)	.129 ± .163 (–.290, .548)	.016 ± .211 (–.541, .574)	.299 ± .263 (–.394, .993)	.40
Diagnostic x-rays of chest or upper body, including mammograms?	2811 / 3163	.130 ± .162 (–.257, .518)	.141 ± .162 (–.276, .558)	.503 ± .573 (–1.01, 2.01)	.110 ± .169 (–.335, .555)	.51
Diagnostic x-rays of the stomach or mid-back?	691 / 3110	.175 ± .164 (–.218, .567)	.178 ± .164 (–.244, .601)	.141 ± .183 (–.342, .625)	.327 ± .370 (–.649, 1.30)	.65
Barium enema?	821 / 3149	.122 ± .162 (–.266, .510)	.123 ± .162 (–.295, .540)	.200 ± .189 (–.298, .698)	–.097 ± .318 (–.937, .743)	.42
Upper GI?	1140 / 3167	.126 ± .162 (–.262, .513)	.122 ± .162 (–.295, .539)	.067 ± .204 (–.471, .604)	.216 ± .267 (–.488, .921)	.66
Intravenous pyelogram?	396 / 3147	.143 ± .162 (–.246, .532)	.146 ± .163 (–.273, .565)	.090 ± .173 (–.366, .545)	.585 ± .482 (–.686, 1.86)	.33
Fluoroscopy of the upper body?	246 / 3151	.140 ± .162 (–.249, .528)	.141 ± .162 (–.278, .559)	.125 ± .169 (–.319, .570)	.335 ± .601 (–1.25, 1.92)	.74
Other nuclear scan?	216 / 3152	.132 ± .162 (–.256, .520)	.134 ± .162 (–.284, .552)	.216 ± .168 (–.228, .660)	–1.15 ± .698 (–2.99, .693)	.049
Dental x-rays that did not usually include a lead shield over the neck area?	1644 / 3181	.133 ± .162 (–.254, .520)	.131 ± .162 (–.285, .547)	.280 ± .231 (–.330, .890)	–.013 ± .227 (–.611, .585)	.36

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

Table IX.P-9. Confounding and Effect Modification by Occupational History: Any Ultrasound-Detected Abnormality

Have You Ever Worked in Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted Estimate	Incl. Confounding Estimate	Including Effect Modification		
				Group 0	Group 1	P
Any metal industry?	238 / 3181	.133 ± .162 (–.254, .520)	.130 ± .162 (–.286, .546)	.114 ± .166 (–.325, .553)	.385 ± .665 (–1.37, 2.14)	.69
Any nuclear facility?	370 / 3181	.133 ± .162 (–.254, .520)	.119 ± .163 (–.300, .539)	.063 ± .180 (–.412, .538)	.370 ± .379 (–.630, 1.37)	.46
Any othe industry or occupation where you may have been exposed to radioactive materials or x-rays?	442 / 3181	.133 ± .162 (–.254, .520)	.132 ± .162 (–.284, .548)	.236 ± .179 (–.237, .709)	–.345 ± .398 (–1.40, .706)	.17
Any of the above industries or occupations?	891 / 3181	.133 ± .162 (–.254, .520)	.138 ± .162 (–.280, .555)	.154 ± .200 (–.374, .683)	.106 ± .276 (–.622, .834)	.89

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

Table IX.P-10 displays the results of analyses of possible confounding or effect modification by smoking variables. There was some evidence that the dose-response coefficient differed between participants with versus without histories of smoking cigarettes (p = 0.034) or any of cigarettes, cigars or pipes (p = 0.024). The estimated dose-response coefficients were greater than zero among nonsmokers, but negative for smokers. However the Bonferroni-adjusted 95% confidence intervals for the smokers' and nonsmokers' estimated coefficients overlapped, including the value of zero in the overlap. In view of the modest significance levels of the effect modification and the large number of comparisons performed in these analyses, these results do not provide compelling evidence of a statistically significant dose-response within the nonsmoking cohort.

Table IX.P-10. Confounding and Effect Modification by Smoking: Any Ultrasound-Detected Abnormality

Have You Ever Smoked Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted Estimate	Incl. Confounding Estimate	Including Effect Modification		
				Group 0	Group 1	P
Cigarettes (unfiltered or filtered)?	1850 / 3173	.140 ± .162 (–.248, .527)	.139 ± .162 (–.278, .556)	.620 ± .281 (–.120, 1.36)	–.109 ± .200 (–.637, .420)	.034
Any of cigarettes, cigar or pipe?	1896 / 3173	.140 ± .162 (–.248, .527)	.139 ± .162 (–.278, .556)	.661 ± .285 (–.092, 1.41)	–.118 ± .199 (–.643, .407)	.024

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

P.2.j. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for the outcome of any thyroid UDA are shown in Figure IX.P-1 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in that figure are calculated at the 98.33% confidence level, i.e., are adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates. While the point estimate of the slope is greater than 0 for 87 of the 100 realizations, the confidence interval includes 0 for all but 1 of the realizations. Also shown in Figure IX.P-1 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean and mean of each participant's 100 dose realizations. In summary, for only one of the 100 realizations of the estimated doses was there a statistically significant dose-response, although for most of the realizations the estimated slope was greater than 0.

Figure IX.P-1. Plot of Estimated Slope and 95% CI by Dose Realization: Any Ultrasound-Detected Abnormality

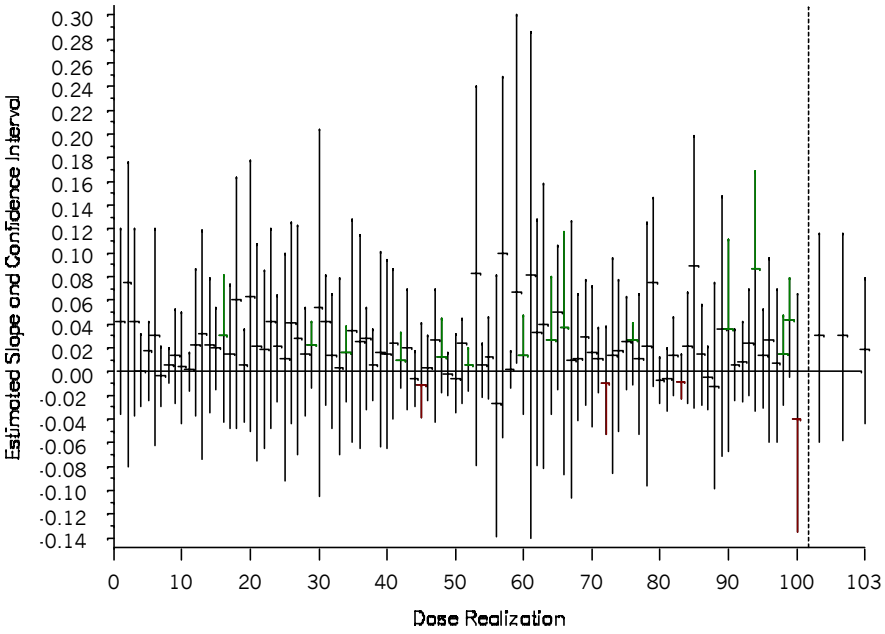
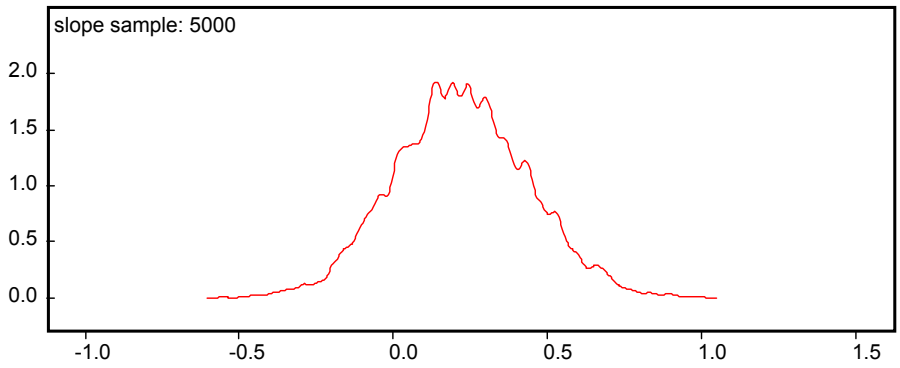


Figure IX.P-2 displays the distribution of the 5000 logistic regression coefficient estimates obtained by the simulation procedures described in section VIII.C.2.b.3 above. It is evident from the figure that most of the estimates were between about –0.5 and 1.0. The estimate was less than or equal to 0 for 759 of the 5000 replications, implying an empirical one-tailed p-value of 0.15. The median estimate was 0.22, and the upper and lower percentiles corresponding to the Bonferroni-adjusted 95% confidence interval were –0.29 and 0.74. These may be compared to the p-value of 0.21 and the estimate of 0.13 with confidence internal (–0.25, 0.52) obtained using the median dose estimates without adjustment for uncertainty. Thus, this method of adjusting the estimated logistic regression coefficient for the uncertainty in the dose estimates did not provide evidence that the prevalence of any thyroid UDA increased significantly with increasing dose.

Figure IX.P-2. Distribution of Simulation Estimates of Logistic Regression Coefficient: Any Ultrasound-Detected Abnormality



P.2.k. Analyses of Numbers of Thyroid UDAs

In the analyses described above, participants were classified according to whether they did or did not have any thyroid UDAs. Additional analyses were performed to investigate whether the number of thyroid UDAs detected in individual participants might increase in relation to estimated thyroid radiation dose. For each living evaluable participant with an HTDS ultrasound examination, the numbers of focal thyroid UDAs with maximum dimension ≥ 5 mm, maximum dimension ≥ 10 mm, and average dimension ≥ 15 mm were counted. These numbers of thyroid UDAs are summarized in Tables IX.P-11 through IX.P-13 below. As shown in Table IX.P-11, study participants had as many as nine thyroid UDAs with maximum dimension ≥ 5 mm, although the majority (60% of the women and 74% of the men) had no such thyroid UDAs. The overall average number of thyroid UDAs of this size was 0.84 per person for women, and 0.47 per person for men.

Table IX.P-11. Number of Ultrasound-Detected Abnormalities of the Thyroid with Maximum Dimension ≥ 5 mm, by Sex and Dose Category

A. Female

Thyroid Radiation Dose (mGy)	Living Evaluable Female No.	Number of Ultrasound-Detected Abnormalities of the Thyroid with Maximum Dimension ≥ 5 mm										
		Avg.	0	1	2	3	4	5	6	7	8	9
OOA*	124	0.94	74 59.7%	26 21.0%	9 7.3%	6 4.8%	2 1.6%	2 1.6%	1 0.8%	2 1.6%	1 0.8%	1 0.8%
< 10	182	0.82	109 59.9%	40 22.0%	13 7.1%	9 4.9%	5 2.7%	2 1.1%	2 1.1%	2 1.1%	0	0
10-49	318	0.83	192 60.4%	66 20.8%	22 6.9%	15 4.7%	13 4.1%	4 1.3%	5 1.6%	1 0.3%	0	0
50-99	311	0.81	193 62.1%	59 19.0%	26 8.4%	15 4.8%	6 1.9%	4 1.3%	6 1.9%	1 0.3%	1 0.3%	0
100-149	220	0.89	123 55.9%	55 25.0%	19 8.6%	8 3.6%	9 4.1%	2 0.9%	0	1 0.5%	2 0.9%	1 0.5%
150-199	125	0.83	76 60.8%	27 21.6%	9 7.2%	2 1.6%	6 4.8%	2 1.6%	2 1.6%	1 0.8%	0	0
200-299	137	0.86	78 56.9%	29 21.2%	17 12.4%	5 3.6%	4 2.9%	1 0.7%	2 1.5%	1 0.7%	0	0
300-399	143	0.81	88 61.5%	30 21.0%	9 6.3%	10 7.0%	1 0.7%	1 0.7%	1 0.7%	1 0.7%	2 1.4%	0
400-999	171	0.82	98 57.3%	38 22.2%	19 11.1%	8 4.7%	5 2.9%	1 0.6%	1 0.6%	0	0	1 0.6%
1000+	7	0.14	6 85.7%	1 14.3%	0	0	0	0	0	0	0	0
Total	1738	0.84	1037 59.7%	371 21.4%	143 8.2%	78 4.5%	51 2.9%	19 1.1%	20 1.2%	10 0.6%	6 0.4%	3 0.2%

*OOA = Out of Area

Table IX.P-11. Number of Ultrasound-Detected Abnormalities of the Thyroid with Maximum Dimension ≥ 5 mm, by Sex and Dose Category (continued)

B. Male

Thyroid Radiation Dose (mGy)	Living Evaluable Male No.	Number of Ultrasound-Detected Abnormalities of the Thyroid with Maximum Dimension ≥ 5 mm										
		Avg.	0	1	2	3	4	5	6	7	8	9
OOA	124	0.56	85 68.6%	22 17.7%	11 8.9%	3 2.4%	1 0.8%	0	1 0.8%	1 0.8%	0	0
< 10	185	0.49	133 71.9%	30 16.2%	12 6.5%	3 1.6%	7 3.8%	0	0	0	0	0
10-49	314	0.44	236 75.2%	50 15.9%	13 4.1%	5 1.6%	6 1.9%	3 1.0%	0	0	0	1 0.3%
50-99	310	0.42	238 76.8%	46 14.8%	12 3.9%	6 1.9%	3 1.0%	1 0.3%	3 1.0%	0	1 0.3%	0
100-149	171	0.42	131 76.6%	25 14.6%	8 4.7%	1 0.6%	3 1.8%	3 1.8%	0	0	0	0
150-199	109	0.53	75 68.8%	21 19.3%	5 4.6%	6 5.5%	1 0.9%	1 0.9%	0	0	0	0
200-299	148	0.72	95 64.2%	30 20.3%	14 9.5%	2 1.4%	2 1.4%	0	3 2.0%	1 0.7%	0	1 0.7%
300-399	160	0.47	120 75.0%	25 15.6%	6 3.8%	4 2.5%	2 1.3%	1 0.6%	1 0.6%	1 0.6%	0	0
400-999	153	0.22	126 82.4%	22 14.4%	4 2.6%	0	1 0.7%	0	0	0	0	0
1000+	17	0.82	11 64.7%	3 17.7%	1 5.9%	1 5.9%	0	0	1 5.9%	0	0	0
Total	1691	0.47	1250 73.9%	274 16.2%	86 5.1%	31 1.8%	26 1.5%	9 0.5%	9 0.5%	3 0.2%	1 0.1%	2 0.1%

*OOA = Out of Area

Focal thyroid UDAs of larger sizes were necessarily less frequent. As shown in Table IX.P-12, participants had as many as eight focal thyroid UDAs with maximum dimension ≥ 10 mm. Again, the majority (78% of women and 86% of men) had no such thyroid UDAs, and the overall average number of thyroid UDAs of this size was 0.34 for women and 0.19 for men.

Table IX.P-12. Number of Ultrasound-Detected Abnormalities of the Thyroid with Maximum Dimension ≥ 10 mm, by Sex and Dose Category

A. Female

Thyroid Radiation Dose (mGy)	Living Evaluable Female No.	Number of Ultrasound-Detected Abnormalities of the Thyroid with Maximum Dimension ≥ 10 mm								
		Avg.	0	1	2	3	4	5	6	8
OOA	124	0.34	98 79.0%	17 13.7%	6 4.8%	0	2 1.6%	1 0.8%	0	0
< 10	182	0.36	139 76.4%	29 15.9%	8 4.4%	4 2.2%	2 1.1%	0	0	0
10-49	318	0.33	245 77.0%	51 16.0%	14 4.4%	6 1.9%	2 0.6%	0	0	0
50-99	311	0.30	246 79.1%	48 15.4%	12 3.9%	2 0.6%	1 0.3%	2 0.6%	0	0
100-149	220	0.40	167 75.9%	33 15.0%	13 5.9%	4 1.8%	2 0.9%	0	0	1 0.5%
150-199	125	0.42	93 74.4%	22 17.6%	5 4.0%	1 0.80	3 2.4%	1 0.8%	0	0
200-299	137	0.26	111 81.0%	20 14.6%	4 2.9%	1 0.7%	1 0.7%	0	0	0
300-399	143	0.31	117 81.8%	15 10.5%	9 6.3%	0	1 0.7%	0	0	1 0.7%
400-999	171	0.43	125 73.1%	30 17.5%	11 6.4%	1 0.6%	2 1.2%	1 0.6%	1 0.6%	0
1000+	7	0.00	7 100%	0	0	0	0	0	0	0
Total	1738	0.34	1348 77.6%	265 15.3%	82 4.7%	19 1.1%	16 0.9%	5 0.3%	1 0.1%	2 0.1%

*OOA = Out of Area

Table IX.P-12. Number of Ultrasound-Detected Abnormalities of the Thyroid with Maximum Dimension ≥ 10 mm, by Sex and Dose Category (continued)

B. Male

Thyroid Radiation Dose (mGy)	Living Evaluable Male No.	Number of Ultrasound-Detected Abnormalities of the Thyroid with Maximum Dimension ≥ 10 mm								
		Avg.	0	1	2	3	4	5	6	8
OOA	124	0.16	107 86.3%	14 11.3%	3 2.4%	0	0	0	0	0
< 10	185	0.20	156 84.3%	22 11.9%	6 3.2%	1 0.5%	0	0	0	0
10-49	314	0.18	267 85.0%	40 12.7%	5 1.6%	2 0.6%	0	0	0	0
50-99	310	0.19	270 87.1%	31 10.0%	4 1.3%	4 1.3%	0	0	0	1 0.3%
100-149	171	0.17	148 86.6%	19 11.1%	3 1.8%	0	1 0.6%	0	0	0
150-199	109	0.18	93 85.3%	14 12.8%	1 0.9%	0	1 0.9%	0	0	0
200-299	148	0.31	119 80.4%	19 12.8%	7 4.7%	1 0.7%	1 0.7%	0	1 0.7%	0
300-399	160	0.19	142 88.8%	13 8.1%	1 0.6%	2 1.3%	1 0.6%	1 0.7%	0	0
400-999	153	0.07	143 93.5%	9 5.9%	1 0.7%	0	0	0	0	0
1000+	17	0.35	14 82.4%	2 11.8%	0	0	1 5.9%	0	0	0
Total	1691	0.19	1459 86.3%	183 10.8%	31 1.8%	10 0.6%	5 0.3%	1 0.1%	1 0.1%	1 0.1%

*OOA = Out of Area

As shown in Table IX.P-13 participants had as many as six focal thyroid UDAs with average dimension ≥ 15 mm. Again, the majority (94% of women and 96% of men) had no such thyroid UDAs, and the overall average number of thyroid UDAs of this size was 0.07 for women and 0.05 for men.

Table IX.P-13. Number of Ultrasound-Detected Abnormalities of the Thyroid with Average Dimension ≥ 15 mm, by Sex and Dose Category

A. Female

Thyroid Radiation Dose (mGy)	Living Evaluable Female No.	Number of Ultrasound-Detected Abnormalities of the Thyroid with Average Dimension ≥ 15 mm					
		Avg.	0	1	2	3	6
OOA	124	0.03	120 96.8%	4 3.2%	0	0	0
< 10	182	0.06	172 94.5%	9 5.0%	1 0.6%	0	0
10-49	318	0.06	302 95.0%	13 4.1%	3 0.9%	0	0
50-99	311	0.06	295 94.9%	14 4.5%	1 0.3%	1 0.3%	0
100-149	220	0.08	207 94.1%	10 4.6%	2 0.9%	1 0.5%	0
150-199	125	0.12	111 88.8%	13 10.4%	1 0.8%	0	0
200-299	137	0.05	131 95.6%	5 3.7%	1 0.7%	0	0
300-399	143	0.08	132 92.3%	10 7.0%	1 0.7%	0	0
400-999	171	0.12	156 91.2%	11 6.4%	3 1.8%	1 0.6%	0
1000+	7	0.00	7 100%	0	0	0	0
Total	1738	0.07	1633 94.0%	89 5.1%	13 0.8%	3 0.2%	0

*OOA = Out of Area

Table IX.P-13. Number of Ultrasound-Detected Abnormalities of the Thyroid with Average Dimension ≥ 15 mm, by Sex and Dose Category (continued)

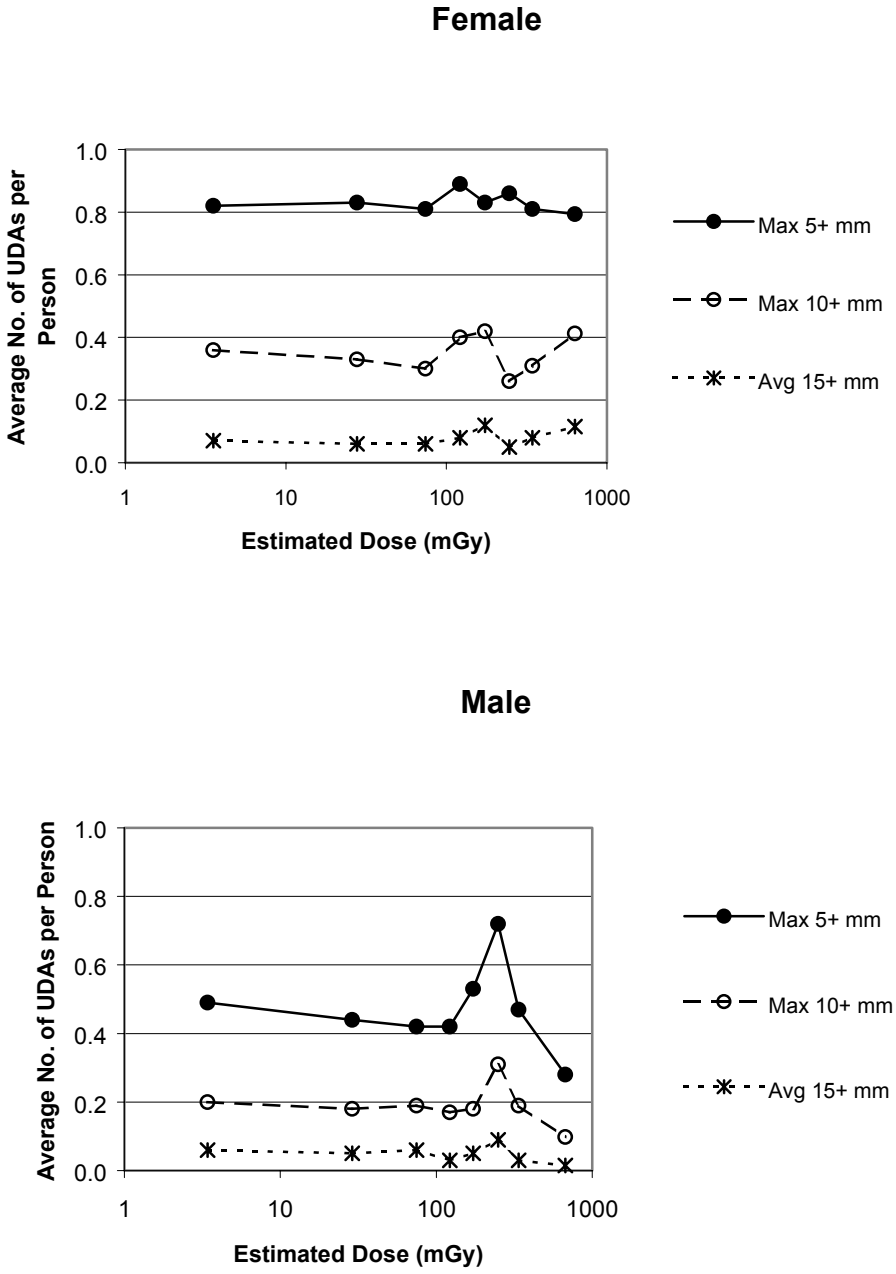
B. Male

Thyroid Radiation Dose (mGy)	Living Evaluable Male No.	Number of Ultrasound-Detected Abnormalities of the Thyroid with Average Dimension ≥ 15 mm					
		Avg.	0	1	2	3	6
OOA	124	0.02	122 98.4%	1 0.8%	1 0.8%	0	0
< 10	185	0.06	176 95.1%	7 3.8%	1 0.5%	1 0.5%	0
10-49	314	0.05	300 95.5%	13 4.1%	1 0.32%	0	0
50-99	310	0.06	298 96.1%	10 3.2%	1 0.3%	0	1 0.3%
100-149	171	0.03	167 97.7%	3 1.8%	1 0.6%	0	0
150-199	109	0.05	104 95.4%	5 4.6%	0	0	0
200-299	148	0.09	139 93.9%	5 3.4%	3 2.0%	1 0.7%	0
300-399	160	0.03	156 97.5%	4 2.5%	0	0	0
400-999	153	0.01	152 99.4%	1 0.7%	0	0	0
1000+	17	0.06	16 94.1%	1 5.9%	0	0	0
Total	1691	0.05	1630 96.4%	50 3.0%	8 0.5%	2 0.1%	1 0.1%

*OOA = Out of Area

Figure IX.P-3 below shows how the average numbers of thyroid UDAs per person, for each of the three size criteria, varied in relation to sex and estimated dose for living evaluable in-area participants. Due to the small number of participants in the 1000+ mGy dose category, it is combined with the 400-999 mGy category in Figure IX.P-3.

Figure IX.P-3. Average Number of Thyroid UDAs per Person, by Sex, Dose Category, and UDA Size



Results of fitting sex-stratified Poisson regression models for the relationship between estimated thyroid radiation dose and number of focal thyroid UDAs are summarized in Table IX.P-14 below. In this table, the estimated dose-response parameter represents the multiplicative change per Gy in the average number of thyroid UDAs per person. For none of these three size criteria did the average number of thyroid UDAs per person increase significantly with increasing estimated dose. For example, for focal thyroid UDAs with maximum dimension ≥ 5 mm, the average number of such thyroid UDAs per person decreased by an estimated factor of 1– 0.92 = 0.08 or 8% for each increase of 1 Gy in the estimated dose.

Consequently the average number of such thyroid UDAs per person did not increase significantly with estimated dose (p = 0.80). The Bonferroni-adjusted 95% confidence interval for the dose-response parameter ranged from 0.72 to 1.17, encompassing a range from a 28% decrease to a 17% increase per Gy.

Table IX.P-14. Poisson Regression Analyses of Numbers of Thyroid UDAs

Size Criterion For Focal Thyroid UDAs	Estimated Background Averages		Estimated Dose-response Parameter (per Gy)	Statistical Significance of Dose-response (one-tailed p-value)
	Female	Male		
Max ≥ 5 mm	0.84 (0.78, 0.91)	0.47 (0.42, 0.51)	0.92 (0.72, 1.17)	0.80
Max ≥ 10 mm	0.34 (0.30, 0.39)	0.19 (0.16, 0.22)	1.01 (0.70, 1.46)	0.48
Avg ≥ 15 mm	0.07 (0.06, 0.10)	0.05 (0.03, 0.06)	1.05 (0.50, 2.23)	0.43

For focal UDAs with average diameter ≥ 15 mm, the average number of such UDAs per person increased by an estimated factor of 1.05 – 1 = 0.05 or 5% for each increase of 1 Gy in the estimated dose (p = 0.43). The Bonferroni-adjusted 95% confidence interval for the dose-response parameter encompassed a range from a 50% decrease to a 123% increase per Gy.

P.3. Palpable Ultrasound-Detected Abnormalities of the Thyroid

Of the 3429 living evaluable participants whose thyroids were visible on the HTDS ultrasound, 224 (6.5%) had palpable ultrasound-detected abnormalities (Table IX.P-15). The ultrasound-detected thyroid abnormalities were based only on the HTDS evaluation.

Table IX.P-15. Proportion of Participants with HTDS Ultrasound Findings of Palpable Thyroid UDAs, by Sex

	Female		Male		Total	
Ultrasound Finding	No.	%	No.	%	No.	%
Thyroid Gland visible on ultrasound	1738	100.0	1691	100.0	3429	100.0
Palpable thyroid UDAs	154	8.9	70	4.1	224	6.5

P.3.a. Primary Analysis

The number and proportion of living evaluable participants with palpable thyroid UDAs is shown by sex, in-area status, and dose group in Table IX.P-16.

Table IX.P-16. Palpable Ultrasound-Detected Abnormalities by Sex and Estimated Dose

Thyroid Radiation Dose (mGy)	Female			Male		
	L.E. with Ultrasound	Palpable Thyroid UDA		L.E. with Ultrasound	Palpable Thyroid UDA	
	No.	No.	%	No.	No.	%
Out of Area	124	14	11.3	124	6	4.8
< 10	182	17	9.3	185	8	4.3
10-49	318	28	8.8	314	17	5.4
50-99	311	27	8.7	310	15	4.8
100-149	220	18	8.2	171	5	2.9
150-199	125	12	9.6	109	4	3.7
200-299	137	11	8.0	148	10	6.8
300-399	143	9	6.3	160	3	1.9
400-999	171	18	10.5	153	2	1.3
1000+	7	0	--	17	0	--
Total	1738	154	8.9	1691	70	4.1

L.E. = living evaluable participants

Of the 224 living evaluable participants with a palpable thyroid UDA, 20 were out-of-area participants. Parameter estimates for the linear dose-response model based on the 3181 in-area participants with ultrasound results are shown in Table IX.P-17. Based on maximum likelihood analysis of the sex-stratified linear probability model, the risk of having palpable thyroid UDA did not increase significantly with estimated dose ($p = 0.95$), with a negative estimated slope B of –0.018 per Gy (Table IX.P-17, row 1). The Bonferroni-adjusted lower 95% confidence limit was not estimated due to the magnitude of the negative slope estimate, however the upper confidence limit was 0.015 per Gy. Estimation by least squares using either the ungrouped or grouped data gave nearly identical results (Table IX.P-17, rows 2 and 3).

Table IX.P-17. Dose-Response Results for Diagnoses of Palpable Thyroid UDA

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
1.	Primary definition	Linear	Primary	None	MLE	.090 ± .008 (.070, .110)	.043 ± .006 (.029, .057)	–.018 ± .023 (NE, .015)	0.95
2.	Primary definition	Linear	Primary	None	LSU	.090 ± .007 (.074, .107)	.044 ± .007 (.027, .061)	–.020 ± .019 (–.066, .027)	0.85
3.	Primary definition	Linear	Primary	None	LSG	.091 ± .007 (.074, .109)	.046 ± .007 (.028, .063)	–.027 ± .022 (–.080, .026)	0.89
4.	Primary definition	LQ	Primary	None	MLE	.090 ± .008 (.072, .109)	.045 ± .008 (.026, .064)	Lin: –.022 ± .033 (–.103, .060) Quad: .002 ± .022 (–.053, .056)	Quad: 0.94

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.P-17. Dose-Response Results for Diagnoses of Palpable Thyroid UDA (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
5.	Primary definition	Logistic	Primary	None	MLE	.092 (.073, .115)	.043 (.031, .060)	–.38 ± .37 (–1.27, .51)	0.86
6.	Primary definition	Linear	Primary	Exclude dose > 1000 mGy	MLE	.093 ± .008 (.073, .113)	.045 ± .006 (.031, .060)	–.030 ± .022 (< –.049, >.028)	0.90
7.	Primary definition	Linear	Primary	Exclude dose > 400 mGy	MLE	.091 ± .009 (.070, .112)	.051 ± .008 (.032, .069)	–.053 ± .040 (–.142, .049)	0.90
8.	Primary definition	Linear	Primary	Exclude OK and F/S geostrata	MLE	.088 ± .009 (.067, .109)	.041 ± .006 (.026, .056)	–.017 ± .023 (NE, .021)	0.92

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.P-17. Dose-Response Results for Diagnoses of Palpable Thyroid UDA (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
9.	Primary definition	Linear	Alt. #1	None	MLE	.090 ± .008 (.071, .109)	.044 ± .006 (.029, .059)	-.018 ± .020 (NE, .011)	0.96
10.	Primary definition	Linear	Alt. #2	None	MLE	.090 ± .008 (.070, .110)	.043 ± .006 (.029, .058)	-.019 ± .023 (NE, .003)	0.99
11.	Primary definition	Linear	Primary	Include OOA (scoping analysis #1)	MLE	.092 ± .008 (.073, .111)	.043 ± .006 (.030, .057)	-.018 ± .023 (NE, .014)	0.95
12.	Primary definition	Linear	Primary	Include OOA (scoping analysis #2)	MLE	.092 ± .008 (.073, .111)	.043 ± .006 (.030, .057)	-.018 ± .023 (NE, .012)	0.96

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

P.3.b. Alternative Dose-Response Functions

As shown in row 4 of Table IX.P-17, the estimated regression coefficient for the dose-squared term in the linear-quadratic dose-response model [5] was 0.002 with Bonferroni-adjusted 95% confidence interval ranging from –0.053 to 0.056. Thus the addition of a quadratic term did not significantly improve the fit of the model (p = 0.94).

The regression parameter for the effect of dose in the sex-stratified logistic regression model [2] was estimated as –0.38 with Bonferroni-adjusted 95% confidence interval ranging from –1.27 to 0.51 (Table IX.P-17, row 5). Thus there was no evidence from the logistic regression model that prevalence of palpable thyroid UDAs increased with increasing dose (p = 0.86).

P.3.c. Effect of Excluding Participants in High Dose Categories

In the analyses excluding participants with estimated dose > 1000 mGy, the estimated slope B was negative (–0.030 per Gy, with Bonferroni-adjusted 95% CI ranging from less than –0.049 to greater than 0.028 per Gy), providing no evidence that the prevalence of palpable thyroid UDA increased with increasing dose (p = 0.90; Table IX.P-17, row 6). When participants with estimated dose > 400 mGy were excluded, the estimated slope B was again less than zero (–0.053 per Gy, with Bonferroni-adjusted 95% CI ranging from –0.142 to 0.049 per Gy), again providing no evidence that the prevalence of palpable thyroid UDA increased with increasing dose (p = 0.90; Table IX.P-17, row 7).

P.3.d. Effect of Excluding Okanogan and Ferry/Stevens Geostrata

In the analyses excluding participants from the Okanogan and Ferry/Stevens Geostrata, the estimated slope B was negative, –0.017 per Gy, providing no evidence that the prevalence of palpable thyroid UDAs increased with increasing dose (p = 0.92; Table IX.P-17, row 8). The Bonferroni-adjusted lower 95% confidence limit was not estimated due to the magnitude of the negative slope estimate, however the upper confidence limit was 0.021 per Gy.

P.3.e. Analysis of Palpable Thyroid UDAs in Relation to Alternative Dose Estimates

For both alternative dose estimates the results were virtually the same as the primary analysis, providing no evidence that the prevalence of palpable thyroid UDAs increased with increasing dose (Table IX.P-17, rows 9 and 10).

P.3.f. Scoping Analysis Regarding Out-of-Area Participants

See section VIII.C.1.a.3 for a description of the scoping analyses that were performed to assess the possible impact of the 249 out-of-area participants. For neither of the two scoping analyses was there any evidence that the prevalence of palpable thyroid UDAs increased with increasing dose (p = 0.95 and p = 0.96 for the first and second scoping analyses, respectively; Table IX.P-17, rows 11 and 12).

P.3.g. Analysis of Palpable Thyroid UDAs In Relation to Alternative Representations of Exposure

In the analyses by geostratum and by dichotomous exposure variable, the sex and age-adjusted comparisons of prevalence were performed as described in section VIII.C.2.a.2.

P.3.g.1. Analysis By Geostratum

As shown in Table IX.P-18, among the 3429 living evaluable in area or out-of-area participants with ultrasound results, the proportions with palpable UDAs ranged from 9/68 (13.2% in the Ferry/Stevens Counties geostratum) to 9/177 (5.1%, Richland) for women, and from 5/63 (7.9%, Okanogan County) to 13/501 (2.6%, Pasco/Kennewick) for men ($p = 0.051$ for heterogeneity among the nine geostrata). In particular the percentages with palpable UDAs were somewhat higher in the Okanogan and Ferry/Stevens geostrata (12.6% for women, 6.8% for men) than in the remaining geostrata (8.5% and 3.9%, respectively; $p = 0.0086$). Since it was likely that participants in the Okanogan and Ferry/Stevens geostrata tended to have lower thyroid doses from Hanford's ^{131}I than those in other geostrata, it does not appear that these differences can be attributed to an effect of Hanford's ^{131}I.).

Table IX.P-18. Palpable Ultrasound-Detected Abnormalities, by Geostratum and Sex

	Female			Male			Total		
Geostratum	No.	Cases	%	No.	Cases	%	No.	Cases	%
Richland	177	9	5.1	172	5	2.9	349	14	4.0
Pasco/Kennewick	505	40	7.9	501	13	2.6	1006	53	5.3
Benton County	375	35	9.3	358	20	5.6	733	55	7.5
Franklin County	73	7	9.6	76	2	2.6	149	9	6.0
Adams County	165	14	8.5	156	10	6.4	321	24	7.5
Walla Walla (city)	131	14	10.7	131	5	3.8	262	19	7.3
Walla Walla County	169	17	10.1	164	6	3.7	333	23	6.9
Okanogan County	75	9	12.0	63	5	7.9	138	14	10.1
Ferry/Stevens Counties	68	9	13.2	70	4	5.7	138	13	9.4
Total	1738	154	8.9	1691	70	4.1	3429	224	6.5

P.3.g.2. Analysis by Dichotomous Exposure Variable

Ninety-five (7.6%) of the 1255 participants in these analyses had palpable thyroid UDAs, including 43/580 (7.4%) in the high exposure group and 52/675 (7.7%) in the low exposure group (Table IX.P-19). Thus the proportion of participants with palpable thyroid UDAs was not significantly elevated in the high exposure group ($p = 0.67$).

Table IX.P-19. Palpable Ultrasound-Detected Abnormalities, by Exposure Group and Sex

	Female			Male			Total		
Exposure Group	No.	Cases	%	No.	Cases	%	No.	Cases	%
Low	350	36	10.3	325	16	4.9	675	52	7.7
High	298	32	10.7	282	11	3.9	580	43	7.4
Total	648	68	10.5	607	27	4.4	1255	95	7.6

P.3.h. Confounding and Effect Modification

As described in section VIII above, additional sex-stratified logistic regression models were investigated to examine the possibility that the primary dose-response results for palpable thyroid UDA might be influenced by confounding, and to search for factors that might modify a radiation dose-response. Table IX.P-20 displays results for models including sex, age at first exposure to Hanford ^{131}I (prenatal, or < 180 days), age at HTDS examination, estimated dose from the NTS, history of any cancer other than thyroid, and HTDS interview type. Note that sex was not analyzed as a possible confounder since its effect was already adjusted for in the sex-stratified model. It is evident from Table IX.P-20 that the model was not significantly improved by adjusting for any of the other factors as a potential confounder: none produced a significantly better fit to the data. Since the estimated slope was virtually unaffected by such adjustments, it does not appear that omitting these factors introduces any important bias in the dose-response results.

There is no evidence of any statistically significant effect modification by any of the covariates in Table IX.P-20, with one possible exception. The dose-response was higher for the 1567 males (0.198) than for the 1614 females (0.067). The statistical significance of this difference must be interpreted with caution due to the large number of such comparisons that were performed. Moreover, neither males nor females had a significantly positive dose-response. Therefore, it does not appear that any of the covariates in Table IX.P-20 were significant effect modifiers for the outcome of palpable ultrasound-detected thyroid abnormalities.

Table IX.P-20. Confounding and Effect Modification by Sex, Age at Exposure or HTDS Examination, Estimated Thyroid Dose from NTS, History of Any Cancer Other Than Thyroid, and Interview Type: Palpable Thyroid UDAs

		Estimated Dose-Response Coefficient (per Gy)				
Covariate (0=No, 1=Yes)	Yes / Total	Unadjusted Estimate	Incl. Confounding Estimate	Including Effect Modification Group 0	Including Effect Modification Group 1	P
Female?	1614 / 3181	-.382 ± .373 (–1.27, .510)	Not Applicable	.198 ± .226 (–4.28, .252)	.067 ± .228 (–.847, 1.07)	.019
Prenatal exposure?	1031 / 3181	-.382 ± .373 (–1.27, .510)	–.451 ± .381 (–1.43, .531)	–.441 ± .439 (–1.60, .719)	-.481 ± .763 (–2.49, 1.53)	.96
1st exposure before age 180 days?	1474 / 3181	-.382 ± .373 (–1.27, .510)	–.421 ± .383 (–1.41, .567)	–.167 ± .563 (–1.65, 1.32)	–.638 ± .552 (–2.09, .818)	.55
Age at exam > 50?	1993 / 3181	-.382 ± .373 (–1.27, .510)	–.413 ± .381 (–1.39, .568)	–.578 ± .738 (–2.53, 1.37)	–.348 ± .446 (–1.53, .829)	.79
NTS I-131 dose > 5.3 mGy?	1563 / 3179	-.382 ± .373 (–1.27, .511)	–.385 ± .382 (–1.37, .600)	–.243 ± .479 (–1.51, 1.02)	–.615 ± .648 (–2.33, 1.10)	.64
History of any cancer other than thyroid?	248 / 3176	–.381 ± .373 (–1.27, .510)	–.381 ± .375 (–1.35, .585)	–.434 ± .406 (–1.51, .638)	–.072 ± .886 (–2.41, 2.26)	.72
Expanded In-Person Interview?	1205 / 3181	-.382 ± .373 (–1.27, .510)	–.318 ± .376 (–1.29, .651)	.068 ± .515 (–1.29, 1.43)	–.744 ± .604 (–2.34, .849)	.30

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

Tables IX.P-21 and IX.P-22 display similar results from analyses including history of medical or dental x-ray exposure or occupational exposure as potential confounding or effect modifying factors. There is no evidence of any confounding or statistically significant effect modification.

Table IX.P-21. Confounding and Effect Modification by History of Medical and Dental Radiation Exposures: Palpable Thyroid UDAs

Have You Ever Had: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted Estimate	Incl. Confounding. Estimate	Including Effect Modification		
				Group 0	Group 1	P
CAT scan of the upper body?	775 / 3139	−.424 ± .381 (−1.34, .488)	−.407 ± .376 (−1.38, .562)	−.240 ± .383 (−1.25, .772)	−1.59 ± 1.15 (−4.62, 1.43)	.24
Diagnostic x-rays of the head?	1188 / 3145	−.335 ± .370 (−1.22, .550)	−.337 ± .369 (−1.29, .614)	−.285 ± .442 (−1.45, .882)	−.449 ± .665 (−2.20, 1.31)	.84
Diagnostic x-rays of the neck?	960 / 3157	−.362 ± .371 (−1.25, .527)	−.362 ± .373 (−1.32, .599)	−.020 ± .447 (−1.20, 1.16)	−1.05 ± .738 (−3.00, .891)	.21
Diagnostic x-rays of chest or upper body, including mammograms?	2811 / 3163	−.372 ± .372 (−1.26, .519)	−.352 ± .372 (−1.31, .607)	−1.57 ± 1.66 (−5.94, 2.81)	−.272 ± .377 (−1.27, .723)	.42
Diagnostic x-rays of the stomach or mid-back?	691 / 3110	−.411 ± .385 (−1.33, .510)	−.413 ± .385 (−1.40, .578)	−.159 ± .393 (−1.20, .879)	−2.04 ± 1.20 (−5.21, 1.14)	.11
Barium enema?	821 / 3149	−.378 ± .373 (−1.27, .514)	−.375 ± .373 (−1.34, .585)	−.368 ± .427 (−1.50, .759)	−.398 ± .763 (−2.41, 1.61)	.97
Upper GI?	1140 / 3167	−.368 ± .372 (−1.26, .522)	−.367 ± .372 (−1.32, .591)	−.001 ± .411 (−1.09, 1.08)	−1.20 ± .730 (−3.13, .721)	.14
Intravenous pyelogram?	396 / 3147	−.391 ± .376 (−1.29, .508)	−.378 ± .376 (−1.35, .591)	−.266 ± .387 (−1.29, .754)	−1.31 ± 1.23 (−4.55, 1.92)	.40
Fluoroscopy of the upper body?	246 / 3151	−.349 ± .373 (−1.24, .544)	−.352 ± .373 (−1.31, .610)	−.242 ± .374 (−1.23, .745)	−2.27 ± 1.91 (−7.31, 2.76)	.24
Other nuclear scan?	216 / 3152	−.386 ± .375 (−1.28, .511)	−.395 ± .375 (−1.36, .571)	−.307 ± .378 (−1.31, .691)	−1.92 ± 1.85 (−6.80, 2.95)	.35
Dental x-rays that did not usually include a lead shield over the neck area?	1644 / 3181	.382 ± .373 (−1.27, .510)	−.385 ± .374 (−1.35, .578)	−.173 ± .483 (−1.45, 1.10)	−.658 ± .589 (−2.21, .895)	.52

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

Table IX.P-22. Confounding and Effect Modification by Occupational History: Palpable Thyroid UDAs

Have You Ever Worked in Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy) Unadjusted Estimate	Incl. Confounding Estimate	Including Effect Modification Group 0	Group 1	P
Any metal industry?	238 / 3181	.382 ± .373 (–1.27, .510)	–.373 ± .372 (–1.33, .586)	–.379 ± .378 (–1.38, .619)	–.166 ± 2.14 (–5.81, 5.48)	.92
Any nuclear facility?	370 / 3181	.382 ± .373 (–1.27, .510)	–.378 ± .375 (–1.34, .589)	–.212 ± .386 (–1.23, .807)	–1.67 ± 1.27 (–5.01, 1.67)	.24
Any other industry or occupation where you may have been exposed to radioactive materials or x-rays?	442 / 3181	.382 ± .373 (–1.27, .510)	–.402 ± .375 (–1.37, .563)	–.186 ± .379 (–1.18, .813)	–3.30 ± 1.89 (–8.30, 1.69)	.06
Any of the above industries or occupations?	891 / 3181	.382 ± .373 (–1.27, .510)	–.346 ± .372 (–1.30, .612)	–.111 ± .398 (–1.16, .938)	–1.34 ± .957 (–3.86, 1.18)	.21

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

Table IX.P-23 displays the results of analyses of possible confounding or effect modification by smoking variables. There was no evidence that the dose-response was significantly confounded by either smoking variable, or that there was a dose-response that differed significantly according to smoking history.

Table IX.P-23. Confounding and Effect Modification by Smoking: Palpable Thyroid UDAs

Have You Ever Smoked Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy) Unadjusted Estimate	Incl. Confounding Estimate	Including Effect Modification Group 0	Group 1	P
Cigarettes (unfiltered or filtered)?	1850 / 3173	–.371 ± .372 (–1.26, .520)	–.373 ± .372 (–1.33, .585)	–.416 ± .608 (–2.02, 1.19)	–.346 ± .468 (–1.58, .888)	.93
Any of cigarettes, cigar or pipe?	1896 / 3173	–.371 ± .372 (–1.26, .520)	–.372 ± .372 (–1.33, .585)	–.291 ± .607 (–1.89, 1.31)	–.421 ± .474 (–1.67, .829)	.87

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

P.3.i. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for the outcome of palpable thyroid UDA are shown in Figure IX.P-4 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in that figure are calculated at the 98.33% confidence level, i.e., are adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific

background rates. While the point estimate of the slope is greater than 0 for 1 of the 100 realizations, the confidence interval includes 0 for all of the realizations. Also shown in Figure IX.P-4 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean and mean of each participant's 100 dose realizations. In summary, for none of the 100 realizations of the estimated doses was there a statistically significant dose-response, and for all but one realization the estimated slope was less than 0.

Figure IX.P-4. Plot of Estimated Slope and 95% CI by Dose Realization: Palpable Thyroid UDAs

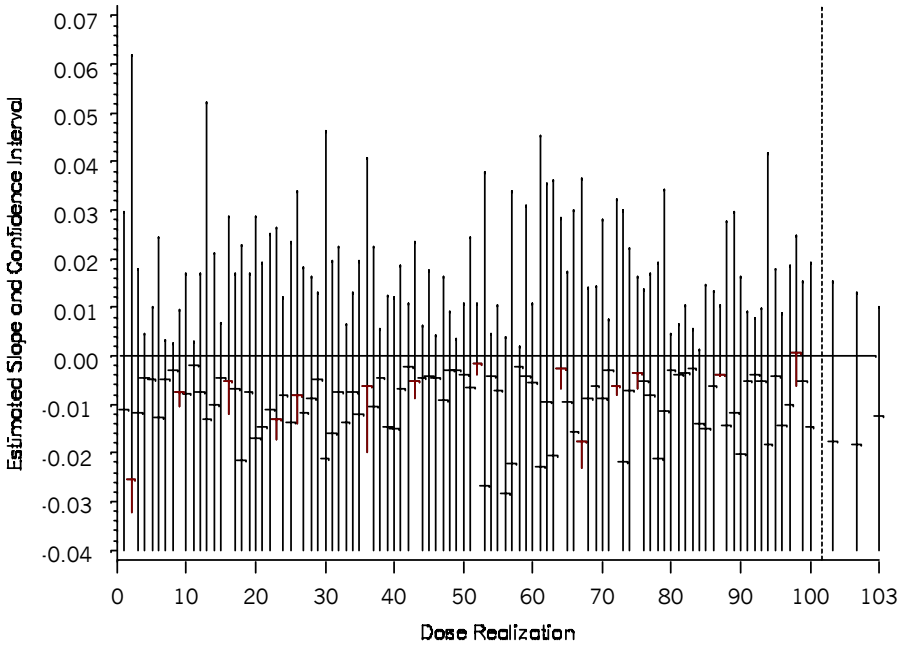
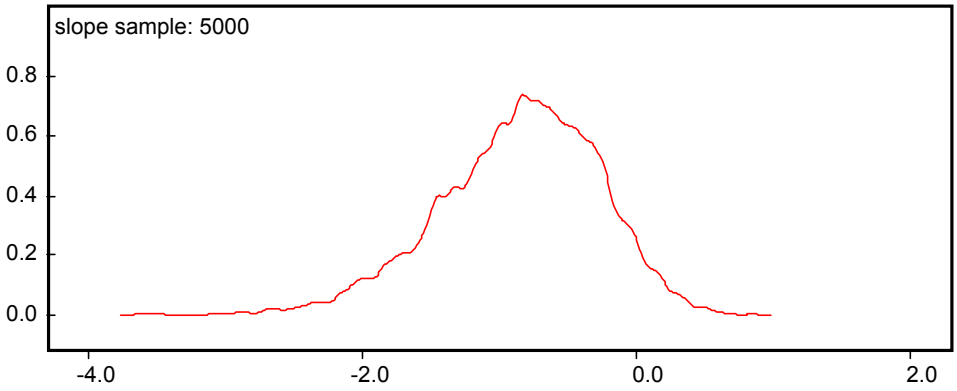


Figure IX.P-5 displays the distribution of the 5000 logistic regression coefficient estimates obtained by the simulation procedures described in section VIII.C.2.b.3 above. It is evident from the figure that most of the estimates were between about –3.0 and 0.3. The estimate was less than or equal to 0 for 4735 of the 5000 replications, implying an empirical one-tailed p-value of 0.95. The median estimate was –0.80, and the upper and lower percentiles corresponding to the Bonferroni-adjusted 95% confidence interval were –2.42 and 0.33. These may be compared to the p-value of 0.21 and the estimate of –0.38 with confidence interval (–1.27, 0.51) obtained using the median dose estimates without adjustment for uncertainty. Thus, this method of adjusting the estimated logistic regression coefficient for the uncertainty in the dose estimates did not provide evidence that the prevalence of palpable thyroid UDA increased with increasing dose.

Figure IX.P-5. Distribution of Simulation Estimates of Logistic Regression Coefficient: Palpable Thyroid UDAs



P.4. Nonpalpable Focal Ultrasound-Detected Abnormalities of the Thyroid

Among the 3429 whose thyroids were visible, 1309 (38.2%) had nonpalpable focal thyroid UDAs. The ultrasound-detected thyroid abnormalities were based only on the HTDS evaluation (Table IX.P-24).

Table IX.P-24. Proportion of Participants with HTDS Ultrasound Findings of Nonpalpable Focal Thyroid UDAs, by Sex

	Female		Male		Total	
Ultrasound Finding	No.	%	No.	%	No.	%
Thyroid gland visible on ultrasound	1738	100.0	1691	100.0	3429	100.0
Nonpalpable focal thyroid UDAs	784	45.1	525	31.0	1309	38.2

P.4.a. Primary Analysis

The proportion with nonpalpable focal thyroid UDAs is shown by sex, in-area status, and dose group in Table IX.P-25.

Table IX.P-25. Nonpalpable Ultrasound-Detected Abnormalities by Sex, and Estimated Dose: Participants with Ultrasound Only

	Female			Male		
Thyroid Radiation Dose	L.E. with Ultrasound	Nonpalpable Focal Thyroid UDA		L.E. with Ultrasound	Nonpalpable Focal Thyroid UDA	
(mGy)	No.	No.	%	No.	No.	%
Out of Area	124	49	39.5	124	43	34.7
< 10	182	81	44.5	185	57	30.8
10-49	318	138	43.4	314	91	29.0
50-99	311	140	45.0	310	84	27.1
100-149	220	109	49.5	171	52	30.4
150-199	125	53	42.4	109	38	34.9
200-299	137	70	51.1	148	56	37.8
300-399	143	65	45.5	160	58	36.3
400-999	171	78	45.6	153	36	23.5
1000+	7	1	14.3	17	10	58.8
Total	1738	784	45.1	1691	525	31.0

L.E. = living evaluable participants

Of the 1309 living evaluable participants with a nonpalpable focal thyroid UDA, 92 were out-of-area participants. Parameter estimates for the linear dose-response model based on the 3181 in-area participants are show in Table IX.P-26 below. The estimated slope B was not significantly greater than zero (0.027 per Gy, with Bonferroni-adjusted 95% confidence interval ranging from -0.061 to 0.115), providing no evidence that the prevalence of nonpalpable thyroid UDAs increased with increasing dose ($p = 0.23$; Table IX.P-26, row 1). Estimation by least squares using either the ungrouped or grouped data gave similar results (Table IX.P-26, rows 2 and 3).

Table IX.P-26. Dose-Response Results for Diagnoses of Nonpalpable Focal Thyroid UDAs

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
1.	Primary definition	Linear	Primary	None	MLE	.451 ± .014 (.417, .484)	.303 ± .013 (.270, .335)	.027 ± .037 (-.061, .115)	0.23
2.	Primary definition	Linear	Primary	None	LSU	.451 ± .014 (.419, .484)	.303 ± .014 (.270, .337)	.024 ± .038 (–.067, .115)	0.27
3.	Primary definition	Linear	Primary	None	LSG	.453 ± .014 (.419, .487)	.305 ± .014 (.271, .340)	.014 ± .044 (–.091, .119)	0.38
4.	Primary definition	LQ	Primary	None	LSU	.442 ± .015 (.405, .479)	.294 ± .015 (.256, .331)	Lin: .111 ± .064 (–.050, .272) Quad: –.072 ± .043 (–.178, .035)	Quad: 0.093

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.P-26. Dose-Response Results for Diagnoses of Nonpalpable Focal Thyroid UDAs (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
5.	Primary definition	Logistic	Primary	None	MLE	.451 (.417, .485)	.304 (.273, .336)	.10 ± .16 (-.29, .49)	0.27
6.	Primary definition	Linear	Primary	Exclude dose > 1000 mGy	MLE	.452 ± .015 (.417, .487)	.300 ± .014 (.266, .333)	.029 ± .048 (<-.085, .145)	0.27
7.	Primary definition	Linear	Primary	Exclude dose > 400 mGy	MLE	.431 ± .016 (.392, .470)	.285 ± .016 (.247, .323)	.228 ± .085 (.026, .431)	0.003
8.	Primary definition	Linear	Primary	Exclude OK and F/S geostrata	MLE	.449 ± .015 (.413, .484)	.295 ± .014 (.261, .329)	.037 ± .038 (-.053, .125)	0.16

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.P-26. Dose-Response Results for Diagnoses of Nonpalpable Focal Thyroid UDAs (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
9.	Primary definition	Linear	Alt. #1	None	MLE	.454 ± .014 (.420, .488)	.306 ± .014 (.273, .339)	.007 ± .037 (-.079, .095)	0.43
10.	Primary definition	Linear	Alt. #2	None	MLE	.446 ± .014 (.412, .480)	.298 ± .014 (.265, .330)	.052 ± .038 (-.038, .142)	0.085
11.	Primary definition	Linear	Primary	Include OOA (scoping analysis #1)	MLE	.446 ± .013 (.414, .478)	.305 ± .013 (.275, .336)	.031 ± .037 (-.056, >.117)	0.20
12.	Primary definition	Linear	Primary	Include OOA (scoping analysis #2)	MLE	.447 ± .013 (.415, .479)	.306 ± .013 (.276, .337)	.025 ± .037 (<-.062, .111)	0.25

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

P.4.b. Alternative Dose-Response Functions

As shown in row 4 of Table IX.P-26, the estimated regression coefficient for the dose-squared term in the linear-quadratic dose-response model [5] was –0.072 with Bonferroni-adjusted 95% confidence interval ranging from –0.178 to 0.035. Thus the addition of a quadratic term did not significantly improve the fit of the model (p = 0.093).

The regression parameter for the effect of dose in the sex-stratified logistic regression model [2] was estimated as 0.10, with Bonferroni-adjusted 95% confidence interval ranging from –0.29 to 0.49 (Table IX.P-26, row 5). Thus there was no evidence from the logistic regression model that prevalence of nonpalpable focal thyroid UDA increased significantly with increasing dose (p = 0.27).

P.4.c. Effect of Excluding Participants in High Dose Categories

As shown in row 7 of Table IX.P-26 above, the estimated slope of the dose-response for nonpalpable focal thyroid UDAs was larger if participants with highest estimated doses were excluded. In particular, when participants with estimated dose > 400 mGy were excluded, the estimated slope B increased from 0.027 to 0.228 per Gy (p = 0.003). Excluding the small number of participants with estimated dose > 1000 mGy had very little effect on the estimated dose-response (Table IX.P-26, row 6).

P.4.d. Effect of Excluding Okanogan and Ferry/Stevens Geostrata

As shown in row 8 of Table IX.P-26, excluding Okanogan and Ferry/Stevens slightly reduced the estimated background rates for both men and women. The reductions were small, because the Okanogan and Ferry/Stevens geostrata account for only 255 (8.0%) of the 3181 in-area living evaluable participants with ultrasound results. As a result of these reductions in the background rates and the fact that Okanogan and Ferry/Stevens geostrata tend to have low doses, the estimated slope changed slightly, from 0.027 to 0.037 per Gy, but remained statistically nonsignificant (p = 0.16).

P.4.e. Analysis of Nonpalpable Focal Thyroid UDAs in Relation to Alternative Dose Estimates

Using the first alternative set of dose estimates, the estimated slope changed slightly, to 0.007 per Gy, with Bonferroni-adjusted 95% confidence limits –0.079 and 0.095 per Gy, which does not represent a statistically significant dose-response for nonpalpable focal thyroid UDAs (p = 0.43; Table IX.P-26, row 9). Similar results were obtained using the second alternative dose estimates, with estimated slope 0.052 per Gy and 95% CI ranging from –0.038 to 0.142 (p = 0.085; Table IX.P-26, row 10).

P.4.f. Scoping Analysis Regarding Out-of-Area Participants

See section VIII.C.1.a.3 for a description of the scoping analyses that were performed to assess the possible impact of the 249 out-of-area participants. The results of both scoping analyses were virtually the same as the primary analysis and provided no evidence that the prevalence of nonpalpable focal thyroid UDA increased with increasing dose (Table IX.P-7). For neither set of scoping analyses was there evidence that the proportion with nonpalpable focal thyroid UDAs increased with increasing dose (p = 0.20 for the first scoping analysis, and p = 0.25 for the second scoping analysis; Table IX.P-26, rows 11 and 12).

P.4.g. Analysis of Nonpalpable Focal Thyroid UDAs in Relation to Alternative Representations of Exposure

In the analyses by geostratum and by dichotomous exposure variable, the sex and age-adjusted comparisons of prevalence were performed as described in section VIII.C.2.a.2.

P.4.g.1. Analysis by Geostratum

The proportions of women with nonpalpable focal thyroid UDAs ranged from 69/131 (52.7%) in the Walla Walla city geostratum to 75/177 (42.4%) in the Richland geostratum (Table IX.P-27). For men they ranged from 32/76 (42.1%) in the Franklin geostratum to 31/164 (18.9%) in the Walla Walla County geostratum. The heterogeneity among the nine geostrata was not statistically significant (p = 0.083). The proportions were somewhat higher in the Okanogan and Ferry/Stevens geostrata (46.2% and 38.3% for women and men, respectively) compared to the other geostrata (45.0% and 30.4%, respectively), also a nonsignificant difference for the heterogeneity between combined geostrata (p = 0.082).

Table IX.P-27. Nonpalpable Focal Ultrasound-Detected Abnormalities, by Geostratum and Sex

	Female			Male			Total		
Geostratum	No.	Cases	%	No.	Cases	%	No.	Cases	%
Richland	177	75	42.4	172	48	27.9	349	123	35.2
Pasco/Kennewick	505	227	45.0	501	156	31.1	1006	383	38.1
Benton County	375	164	43.7	358	114	31.8	733	278	37.9
Franklin County	73	35	47.9	76	32	42.1	149	67	45.0
Adams County	165	74	44.8	156	57	36.5	321	131	40.8
Walla Walla (city)	131	69	52.7	131	36	27.5	262	105	40.1
Walla Walla County	169	74	43.8	164	31	18.9	333	105	31.5
Okanogan County	75	33	44.0	63	26	41.3	138	59	42.8
Ferry/Stevens Counties	68	33	48.5	70	25	35.7	138	58	42.0
Total	1738	784	45.1	1691	525	31.0	3429	1309	38.2

P.4.g.2. Analysis by Dichotomous Exposure Variable

A total of 494 (39.4%) of the 1255 participants in these analyses had nonpalpable focal thyroid UDAs, including 240/580 (41.4%) in the high exposure group and 254/675 (37.6%) in the low exposure group (Table IX.P-28). Based on the logistic regression analysis with adjustment for the effect of sex and age at HTDS examination, the proportion of participants with nonpalpable focal thyroid UDA was not significantly elevated in the high exposure group (p = 0.081).

Table IX.P-28. Nonpalpable Focal Ultrasound-Detected Abnormalities, by Exposure Group and Sex

	Female			Male			Total		
Exposure Group	No.	Cases	%	No.	Cases	%	No.	Cases	%
Low	350	154	44.0	325	100	30.8	675	254	37.6
High	298	149	50.0	282	91	32.3	580	240	41.4
Total	648	303	46.8	607	191	31.5	1255	494	39.4

P.4.h. Confounding and Effect Modification

As described in section VIII above, additional sex-stratified logistic regression models were investigated to examine the possibility that the primary dose-response results might be influenced by confounding, and to search for factors that might modify a radiation dose-response. Table IX.P-29 displays results for models including sex, age at first exposure to Hanford I-131 (prenatal, or < 180 days), age at HTDS examination, estimated dose from the NTS, history of any cancer other than thyroid, and HTDS interview type. Note that sex was not analyzed as a possible confounder since its effect was already adjusted for in the sex-stratified model.

It is evident from Table IX.P-29 that the estimated slope was virtually unaffected by adjustments for possible confounding. Therefore, it does not appear that omitting these factors introduces any important bias in the dose-response results. In addition, the regression coefficients did not differ significantly between the groups defined by any of the covariates, suggesting that none of them were significant modifiers of a radiation dose-response for nonpalpaple UDAs.

Table IX.P-29. Confounding and Effect Modification by Sex, Age at Exposure or HTDS Examination, Estimated Thyroid Dose from NTS, History of Any Cancer Other Than Thyroid, and Interview Type: Nonpalpable Focal Ultrasound-Detected Abnormalities

		Estimated Dose-Response Coefficient (per Gy)				
Covariate (0=No, 1=Yes)	Yes / Total	Unadjusted Estimate	Incl. Confounding. Estimate	Including Effect Modification Group 0	Including Effect Modification Group 1	P
Female?	1614 / 3181	.102 ± .164 (–.290, .494)	Not Applicable	.198 ± .226 (–.208, .949)	.067 ± .228 (–.724, .422)	.11
Prenatal exposure?	1031 / 3181	.102 ± .164 (–.290, .494)	.045 ± .165 (–.381, .471)	.137 ± .189 (–.360, .634)	–.261 ± .349 (–1.18, .659)	.31
1st exposure before age 180 days?	1474 / 3181	.102 ± .164 (–.290, .494)	.092 ± .165 (–.333, .517)	.516 ± .280 (–.224, 1.26)	–.135 ± .210 (–.690, .419)	.062
Age at exam > 50?	1993 / 3181	.102 ± .164 (–.290, .494)	.111 ± .166 (–.317, .538)	–.183 ± .308 (–.996, .630)	.236 ± .198 (–.286, .758)	.25
NTS I-131 dose > 5.3 mGy?	1563 / 3179	.096 ± .164 (–.296, .488)	.063 ± .167 (–.368, .495)	.116 ± .221 (–.467, .698)	–.007 ± .258 (–.689, .675)	.72
History of any cancer other than thyroid?	248 / 3176	.103 ± .164 (–.289, .495)	.107 ± .164 (–.315, .529)	.193 ± .177 (–.275, .661)	–.413 ± .474 (–1.66, .837)	.21
Expanded In-Person Interview?	1205 / 3181	.102 ± .164 (–.290, .494)	.106 ± .167 (–.324, .535)	.401 ± .266 (–.300, 1.10)	–.085 ± .217 (–.658, .488)	.16

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

Tables IX.P-30 and IX.P-31 display similar results from analyses including history of medical or dental x-ray exposure or occupational exposure as potential confounding or effect modifying factors. There was no evidence of confounding, or of clearly significant effect modification, by any of these variables.

Table IX.P-30. Confounding and Effect Modification by History of Medical and Dental Radiation Exposures: Nonpalpable Focal Ultrasound-Detected Abnormalities

Have You Ever Had: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy) Unadjusted Estimate	Incl. Confounding. Estimate	Including Effect Modification Group 0	Including Effect Modification Group 1	P
CAT scan of the upper body?	775 / 3139	.102 ± .164 (–.292, .495)	.102 ± .164 (–.321, .525	–.030 ± .182 (–.509, .450)	.746± .397 (–.302, 1.79)	.08
Diagnostic x-rays of the head?	1188 / 3145	.103 ± .165 (–.292, .498)	.099 ± .165 (–.327, .524)	.242 ± .204 (–.296, .779)	–.181 ± .290 (–.947, .584)	.23
Diagnostic x-rays of the neck?	960 / 3157	.094 ± .165 (–.300, .487)	.100 ± .165 (–.325, .525)	.012 ± .217 (–.560, .585)	.223 ± .257 (–.454, .900)	.53
Diagnostic x-rays of chest or upper body, including mammograms?	2811 / 3163	.098 ± .164 (–.294, .491)	.103 ± .164 (–.320, .526)	.270 ± .590 (–1.29, 1.83)	.089 ± .171 (–.361, .540)	.77
Diagnostic x-rays of the stomach or mid-back?	691 / 3110	.128 ± .166 (–.268, .524)	.131 ± .166 (–.296, .557)	.068 ± .186 (–.423, .559)	.379 ± .370 (–.596, 1.35)	.45
Barium enema?	821 / 3149	.104 ± .164 (–.289, .497)	.105 ± .164 (–.318, .528)	.104 ± .191 (–.400, .608)	.107 ± .321 (–.740, .954)	.99
Upper GI?	1140 / 3167	.100 ± .164 (–.292, .493)	.099 ± .164 (–.323, .521)	.009 ± .208 (–.539, .557)	.252 ± .270 (–.459, .964)	.47
Intravenous pyelogram?	396 / 3147	.102 ± .165 (–.292, .496)	.104 ± .165 (–.320, .528)	.108 ± .175 (–.354, .569)	.072 ± .487 (–1.21, 1.36)	.94
Fluoroscopy of the upper body?	246 / 3151	.118 ± .164 (–.275, .511)	.119 ± .164 (–.305, .542)	.090 ± .171 (–.361, .542)	.477 ± .603 (–1.11, 2.07)	.54
Other nuclear scan?	216 / 3152	.092 ± .164 (–.301, .486)	.093 ± .164 (–.331, .516)	.155 ± .169 (–.291, .602)	–.903 ±.709 (–2.77, .966)	.13
Dental x-rays that did not usually include a lead shield over the neck area?	1644 / 3181	.102 ± .164 (–.290, .494)	.103 ± .164 (–.319, .524)	.212 ± .233 (–.403, .827)	–.003 ± .230 (–.610, .605)	.51

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

Table IX.P-31. Confounding and Effect Modification by Occupational History: Nonpalpable Focal Ultrasound-Detected Abnormalities

Have You Ever Worked in Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy) Unadjusted Estimate	Incl. Confounding Estimate	Including Effect Modification Group 0	Group 1	P
Any metal industry?	238 / 3181	.102 ± .164 (–.290, .494)	.098 ± .164 (–.323, .520)	.071 ± .169 (–.374, .516)	.565 ± .685 (–1.24, 2.37)	.49
Any nuclear facility?	370 / 3181	.102 ± .164 (–.290, .494)	.099 ± .165 (–.326, .525)	–.010 ± .183 (–.493, .474)	.588 ± .385 (–.427, 1.60)	.16
Any other industry or occupation where you may have been exposed to radioactive materials or x-rays?	442 / 3181	.102 ± .164 (–.290, .494)	.101 ± .164 (–.321, .522)	.176 ± .180 (–.300, .652)	–.260 ± .413 (–1.35, .829)	.32
Any of the above industries or occupations?	891 / 3181	.102 ± .164 (–.290, .494)	.110 ± .164 (–.312, .533)	.047 ± .202 (–.487, .580)	.233 ± .280 (–.506, .972)	.59

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

Table IX.P-32 displays the results of analyses of possible confounding or effect modification by smoking variables. There was some evidence that the dose-response coefficient differed between participants with versus without histories of smoking cigarettes (p = 0.033) or any of cigarettes, cigars or pipes (p = 0.019). The estimated dose-response coefficients were greater than zero among nonsmokers, but negative for smokers. However, the Bonferroni-adjusted 95% confidence intervals for the smokers' and nonsmokers' estimated coefficients overlapped, including the value of zero in the overlap. In view of the significance levels of these two tests for effect modification and the large number of comparisons performed in these analyses, these results do not provide compelling evidence of a statistically significant dose-response within the nonsmoking cohort.

Table IX.P-32. Confounding and Effect Modification by Smoking: Nonpalpable Focal Ultrasound-Detected Abnormalities

Have You Ever Smoked Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy) Estimate	Adjusted for Estimate	Including Effect Modification Group 0	Group 1	P
Cigarettes (unfiltered or filtered)?	1850 / 3173	.108 ± .164 (–.284, .500)	.107 ± .164 (–.315, .529)	.589 ± .279 (–.148, 1.33)	–.150 ± .208 (–.699, .398)	.033
Any of cigarettes, cigar or pipe?	1896 / 3173	.108 ± .164 (–.284, .500)	.107 ± .164 (–.315, .529)	.648 ± .284 (–.100, 1.40)	–.172 ± .207 (–.718, .375)	.019

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values reflect the significance of the improved fit for models including confounding or effect modification, relative to the unadjusted model.

P.4.i. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for the outcome of nonpalpable focal thyroid UDA are shown in Figure IX.P-6 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in that figure are calculated at the 98.33% confidence level, i.e., are adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates. While the point estimate of the slope is greater than 0 for 85 of the 100 realizations, the confidence interval includes 0 for all but 1 of the realizations. Also shown in Figure IX.P-6 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean and mean of each participant's 100 dose realizations. In summary, for only one of the 100 realizations of the estimated doses was there a statistically significant dose-response, although for most of the realizations the estimated slope was greater than 0.

Figure IX.P-6. Plot of Estimated Slope and 95% CI by Dose Realization: Nonpalpable Focal *Ultrasound-Detected Abnormalities*

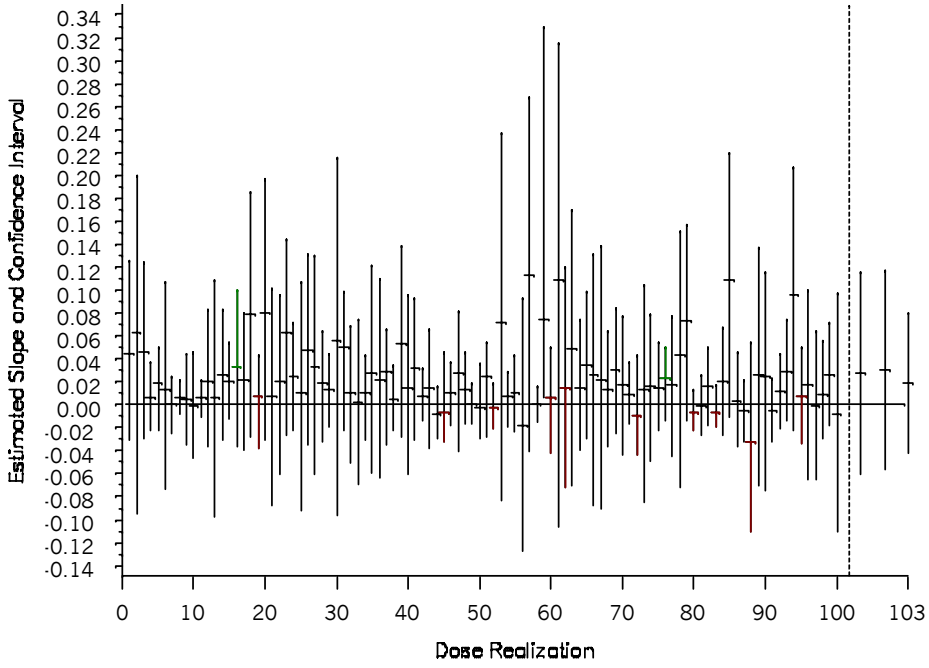
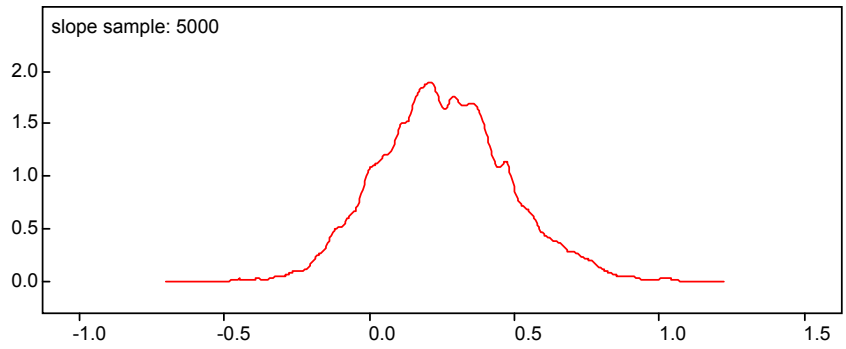


Figure IX.P-7 displays the distribution of the 5000 logistic regression coefficient estimates obtained by the simulation procedures described in section VIII.C.2.b.3 above. It is evident from the figure that most of the estimates were between about –0.5 and 1.0. The estimate was less than or equal to 0 for 604 of the 5000 replications, implying an empirical one-tailed p-value of 0.12. The median estimate was –0.25, and the upper and lower percentiles corresponding to the Bonferroni-adjusted 95% confidence interval were –0.27 and 0.84. These may be compared to the p-value of 0.27 and the estimate of 0.10 with confidence interval (–0.29, 0.49) obtained using the median dose estimates without adjustment for uncertainty. Thus, this method of adjusting the estimated logistic regression coefficient for the uncertainty in the dose estimates did not provide evidence that the prevalence of nonpalpable focal thyroid UDA increased significantly with increasing dose.

Figure IX.P-7. Distribution of Simulation Estimates of Logistic Regression Coefficient: Nonpalpable Focal Ultrasound-Detected Abnormalities



P.5. Diffuse Ultrasound-Detected Abnormalities of the Thyroid

Of the 3429 living evaluable participants whose thyroids were visible, 458 (13.4%) had diffuse ultrasound-detected thyroid abnormalities (Table IX.P-33).

Table IX.P-33. Proportion of Participants with HTDS Ultrasound Findings of Diffuse Thyroid UDAs, by Sex

	Female		Male		Total	
Ultrasound Finding	No.	%	No.	%	No.	%
Thyroid gland visible on ultrasound	1738	100.0	1691	100.0	3429	100.0
Diffuse thyroid UDAs	306	17.6	152	9.0	458	13.4

P.5.a. Primary Analysis

Of the 458 living evaluable participants with a diagnosis of diffuse thyroid UDAs, 30 were out-of-area participants. The proportions with diffuse thyroid UDAs are shown by sex, in-area status and dose group in Table IX.P-34.

Table IX.P-34. Diffuse Ultrasound-Detected Abnormalities by Sex, and Estimated Dose: Participants with Ultrasound Only

Thyroid Radiation Dose (mGy)	Female			Male		
	L.E. with Ultrasound	Diffuse Ultrasound-Detected Abnormality		L.E. with Ultrasound	Diffuse Ultrasound Detected Abnormality	
	No.	No.	%	No.	No.	%
Out of Area	124	18	14.5	124	12	9.7
< 10	182	29	15.9	185	19	10.3
10-49	318	53	16.7	314	21	6.7
50-99	311	47	15.1	310	22	7.1
100-149	220	52	23.6	171	17	9.9
150-199	125	25	20.0	109	9	8.3
200-299	137	24	17.5	148	21	14.2
300-399	143	30	21.0	160	18	11.3
400-999	171	27	15.8	153	11	7.2
1000+	7	1	14.3	17	2	11.8
Total	1738	306	17.6	1691	152	9.0

L.E. = living evaluable participants

Parameter estimates for the linear dose-response model based on the 3181 in-area participants with a visible thyroid are shown in Table IX.P-35 below. The estimated slope B was not significantly greater than zero (0.029 per Gy with Bonferroni-adjusted 95% confidence interval ranging from 0.029 to 0.100) providing no evidence that the proportion with diffuse thyroid UDAs increased with increasing dose ($p = 0.14$; Table IX.P-35, row 1). Estimating by least squares using either the ungrouped or grouped data gave similar results (Table IX.P-35, rows 2 and 3).

Table IX.P-35. Dose-Response Results for Diagnoses of Diffuse Ultrasound Abnormalities

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
1.	Primary definition	Linear	Primary	None	MLE	.174 ± .011 (.148, .199)	.084 ± .009 (.064, .105)	.029 ± .028 (-.029, .100)	0.14
2.	Primary definition	Linear	Primary	None	LSU	.174 ± .010 (.151, .197)	.085 ± .010 (.061, .108)	.026 ± .027 (-.039, .090)	0.17
3.	Primary definition	Linear	Primary	None	LSG	.176 ± .010 (.153, .200)	.087 ± .010 (.063, .111)	.013 ± .031 (-.061, .086)	0.34
4.	Primary definition	LQ	Primary	None	LSU	.174 ± .010 (.148, .200)	.085 ± .011 (.058, .111)	Lin: .027 ± .045 (–.086, .140) Quad: –.001 ± .030 (–.076, .074)	Quad: 0.97

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.P-35. Dose-Response Results for Diagnoses of Diffuse Ultrasound Abnormalities (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
5.	Primary definition	Logistic	Primary	None	MLE	.173 (.149, .201)	.086 (.070, .107)	.21 ± .22 (-.32, .74)	0.17
6.	Primary definition	Linear	Primary	Exclude dose > 1000 mGy	MLE	.173 ± .011 (.147, .199)	.084 ± .009 (.062, .105)	.033 ± .034 (<-.043, >.118)	0.16
7.	Primary definition	Linear	Primary	Exclude dose > 400 mGy	MLE	.164 ± .012 (.136, .193)	.074 ± .010 (.051, .097)	.146 ± .059 (.010, .291)	0.005
8.	Primary definition	Logistic	Primary	Exclude OK and F/S geostrata	MLE	.172 ± .011 (.146, .199)	.075 ± .009 (.054, .097)	.042 ± .029 (-.021, .115)	0.065

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

Table continued on next page

Table IX.P-35. Dose-Response Results for Diagnoses of Diffuse Ultrasound Abnormalities (continued)

Row	Outcome	Dose-Response Model	Dose Estimates	Exclusions / Additional Inclusions	Method of Analysis	Estimated Background Rates		Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
						Female	Male		
9.	Primary definition	Linear	Alt. #1	None	MLE	.176 ± .011 (.151, .202)	.087 ± .009 (.066, .108)	.010 ± .026 (<-.041, .078)	0.34
10.	Primary definition	Linear	Alt. #2	None	MLE	.176 ± .011 (.151, .201)	.087 ± .009 (.066, .108)	.013 ± .026 (-.037, .080)	0.30
11.	Primary definition	Linear	Primary	Include OOA (scoping analysis #1)	MLE	.171 ± .010 (.147, .195)	.085 ± .008 (.065, .104)	.031 ± .027 (-.027, >.101)	0.12
12.	Primary definition	Linear	Primary	Include OOA (scoping analysis #2)	MLE	.172 ± .010 (.148, .196)	.085 ± .008 (.066, .105)	.028 ± .027 (<-.029, >.097)	0.14

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). Standard errors are not given for estimated background rates from logistic regression model.

Abbreviations: MLE = maximum likelihood estimation, LSU = least squares estimation with ungrouped doses, LSG = least squares estimation with grouped doses, LQ = linear-quadratic, Ok = Okanogan, F/S = Ferry/Stevens, OOA = out of area

P.5.b. Alternative Dose-Response Functions

As shown in row 4 of Table IX.P-35, the estimated regression coefficient for the dose-squared term in the linear-quadratic dose-response model [5] was –0.001 with Bonferroni-adjusted 95% confidence interval ranging from –0.076 to 0.074. Thus the addition of a quadratic term did not significantly improve the fit of the model (p = 0.97).

The regression parameter for the effect of dose in the sex-stratified logistic regression model [2] was estimated as 0.21, with Bonferroni-adjusted 95% confidence interval ranging from –0.32 to 0.74 (Table IX.P-35, row 5). Thus there was no evidence from the logistic regression model that prevalence of diffuse thyroid UDA increased significantly with increasing dose (p = 0.17).

P.5.c. Effect of Excluding Participants in High Dose Categories

As shown in row 7 of Table IX.P-35 above, the estimated slope of the dose-response for nonpalpable focal thyroid UDAs was larger if participants with highest estimated doses were excluded. In particular, when participants with estimated dose > 400 mGy were excluded, the estimated slope B increased from 0.029 to 0.146 per Gy (p = 0.005). Excluding the small number of participants with estimated dose > 1000 mGy had very little effect on the estimated dose-response (Table IX.P-35, row 6).

P.5.d. Effect of Excluding Okanogan and Ferry/Stevens Geostrata

The effect of excluding the Okanogan and Ferry/Stevens geostrata was to increase the estimated slope, from 0.029 to 0.042 per Gy. The statistical significance of the dose-response changed from p = 0.14 to p = 0.065 (Table IX.P-35, row 8).

P.5.e. Analysis of Diffuse Thyroid UDAs in Relation to Alternative Dose Estimates

Using the first alternative dose estimates, the estimated slope B was not significantly greater than zero (0.010 per Gy with Bonferroni-adjusted 95% CI ranging from less than -0.041 to 0.078 per Gy), providing no evidence that prevalence increased with increasing dose (p = 0.34; Table IX.P-35, row 9). Similar results were found with the second set of alternative dose estimates, with an estimated slope of (0.013 per Gy with Bonferroni-adjusted 95% CI ranging from -0.037 to 0.080), and no evidence that the proportion with diffuse thyroid UDAs increased with increasing dose (p = 0.30; Table IX.P-35, row 10).

P.5.f. Scoping Analysis Regarding Out-of-Area Participants

See section VIII.C.1.a.3 for a description of the scoping analyses that were performed to assess the possible impact of the 249 out-of-area participants. The results of both scoping analyses were virtually the same as the primary analysis and provided no evidence that the prevalence of diffuse thyroid UDA increased with increasing dose (Table IX.P-35; rows 11 and 12).

P.5.g. Analysis of Diffuse Thyroid UDAs in Relation to Alternative Representations of Exposure

In the analyses by geostratum and by dichotomous exposure variable, the sex and age-adjusted comparisons of prevalence were performed as described in section VIII.C.2.a.2.

P.5.g.1. Analysis by Geostratum

Among women, the proportions with diffuse thyroid UDAs ranged from 28/131 (21.4%) in the Walla Walla City geostratum to 26/177 (14.7%) in the Richland geostratum (Table IX.P-36). For men they ranged from 11/70 (15.7%) in the Ferry/Stevens geostratum to 9/164 (5.5%) in the Walla Walla County geostratum. The heterogeneity among the nine geostrata was not statistically significant (p = 0.60). Among men diffuse thyroid UDAs were rather more common in the Okanogan and Ferry/Stevens geostrata (14.3%) compared to the other geostrata (8.5%). However among women the proportions were nearly identical (16.8% and 17.7%). The difference between the Okanogan and Ferry/Stevens geostrata versus the other geostrata was not statistically significant (p = 0.32).

Table IX.P-36. Diffuse Ultrasound-Detected Abnormalities, by Geostratum and Sex

	Female			Male			Total		
Geostratum	No.	Cases	%	No.	Cases	%	No.	Cases	%
Richland	177	26	14.7	172	13	7.6	349	39	11.2
Pasco/Kennewick	505	96	19.0	501	40	8.0	1006	136	13.5
Benton County	375	62	16.5	358	39	10.9	733	101	13.8
Franklin County	73	12	16.4	76	10	13.2	149	22	14.8
Adams County	165	32	19.4	156	14	9.0	321	46	14.3
Walla Walla (city)	131	28	21.4	131	8	6.1	262	36	13.7
Walla Walla County	169	26	15.4	164	9	5.5	333	35	10.5
Okanogan County	75	13	17.3	63	8	12.7	138	21	15.2
Ferry/Stevens Counties	68	11	16.2	70	11	15.7	138	22	15.9
Total	1738	306	17.6	1691	152	9.0	3429	458	13.4

P.5.g.2. Analysis by Dichotomous Exposure Variable

Of the 1255 participants included in these analyses, 175 (13.9%) had diffuse thyroid UDA (Table IX.P-37). These included 83/580 (14.3%) in the high exposure group, and 92/675 (13.6%) in the low exposure group. Based on the logistic regression analysis with adjustment for the effect of sex and age at HTDS examination, the proportion of participants with diffuse thyroid UDAs was not significantly elevated in the high exposure group (p = 0.25).

Table IX.P-37. Diffuse Ultrasound-Detected Abnormalities, by Exposure Group and Sex

	Female			Male			Total		
Exposure Group	No.	Cases	%	No.	Cases	%	No.	Cases	%
Low	350	57	16.3	325	35	10.8	675	92	13.6
High	298	58	19.5	282	25	8.9	580	83	14.3
Total	648	115	17.7	607	60	9.9	1255	175	13.9

P.5.h. Confounding and Effect Modification

As described in section VIII above, additional sex-stratified logistic regression models were investigated to examine the possibility that the primary dose-response results might be influenced by confounding, and to search for factors that might modify a radiation dose-response. Table IX.P-38 displays results for models including sex, age at first exposure to Hanford I-131 (prenatal, or < 180 days), age at HTDS examination, estimated dose from the NTS, history of any cancer other than thyroid, and HTDS interview type. Note that sex was not analyzed as a possible confounder since its effect was already adjusted for in the sex-stratified model. It is evident from Table IX.P-38 that the model was not significantly improved by adjusting for any of the other factors as a potential confounder: none produced a significantly better fit to the data. Since the estimated slope was virtually unaffected by such adjustments, it does not appear that omitting these factors introduces any important bias in the dose-response results.

Table IX.P-38. Confounding and Effect Modification by Sex, Age at Exposure or HTDS Examination, Estimated Thyroid Dose from NTS, History of Any Cancer Other than Thyroid, and Interview Type: Diffuse Ultrasound-Detected Abnormalities

		Estimated Dose-Response Coefficient (per Gy)				
Covariate (0=No, 1=Yes)	Yes / Total	Unadjusted Estimate	Incl. Confounding Estimate	Including Effect Modification Group 0	Including Effect Modification Group 1	P
Female?	1614 / 3181	.211 ± .220 (–.317, .738)	Not Applicable	.392 ± .343 (–.464, 1.25)	.095 ± .286 (–.621, .810)	.51
Prenatal exposure?	1031 / 3181	.211 ± .220 (–.317, .738)	.188 ± .223 (–.385, .762)	.364 ± .242 (–.274, 1.00)	–.510 ± .534 (–1.92, .898)	.12
1st exposure before age 180 days?	1474 / 3181	.211 ± .220 (–.317, .738)	.211 ± .223 (–.363, .785)	.254 ± .392 (–.779, 1.29)	.190 ± .273 (–.529, .909)	.89
Age at exam > 50?	1993 / 3181	.211 ± .220 (–.317, .738)	.292 ± .220 (–.275, .859)	.479 ± .359 (–.469, 1.43)	.185 ± .282 (–.559, .928)	.52
NTS I-131 dose > 5.3 mGy?	1563 / 3179	.211 ± .220 (–.316, .739)	.179 ± .226 (–.404, .763)	.358 ± .286 (–.396, 1.11)	–.107 ± .390 (–1.14, .922)	.33
History of any cancer other than thyroid ?	248 / 3176	.220 ± .220 (–.307, .746)	.219 ± .220 (–.348, .785)	.213 ± .244 (–.431, .857)	.243 ± .508 (–1.10, 1.58)	.96
Expanded In-Person Interview?	1205 / 3181	.211 ± .220 (–.317, .738)	.232 ± .224 (–.346, .810)	.450 ± .361 (–.503, 1.40)	.097 ± .296 (–.684, .877)	.45

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

Tables IX.P-39 and IX.P-40 display similar results from analyses including history of medical or dental x-ray exposure or occupational exposure as potential confounding or effect modifying factors. There is no evidence of any confounding or statistically significant effect modification.

Table IX.P-39. Confounding and Effect Modification by History of Medical and Dental Radiation Exposures: Diffuse Ultrasound-Detected Abnormalities

		Estimated Dose-Response Coefficient (per Gy)				
Have You Ever Had: (0=No, 1=Yes)	Yes / Total	Unadjusted Estimate	Incl. Confounding Estimate	Including Effect Modification Group 0	Including Effect Modification Group 1	P
CAT scan of the upper body?	775 / 3139	.195 ± .222 (–.228, .727)	.192 ± .223 (–.381, .766)	.155 ± .247 (–.496, .806)	.376± .539 (–1.04, 1.80)	.71
Diagnostic x-rays of the head?	1188 / 3145	.215 ± .222 (–.316, .745)	.217 ± .221 (–.353, .787)	.242 ± .272 (–.477, .961)	.169 ± .380 (–.834, 1.17)	.88
Diagnostic x-rays of the neck?	960 / 3157	.229 ± .221 (–.299, .757)	.249 ± .221 (–.320, .819)	–.152 ± .325 (–1.01, .705)	.662 ± .304 (–.141, 1.46)	.066
Diagnostic x-rays of chest or upper body, including mammograms?	2811 / 3163	.218 ± .220 (–.309, .745)	.241 ± .220 (–.326, .808)	.563 ± .800 (–1.55, 2.67)	.216 ± .230 (–.390, .822)	.68
Diagnostic x-rays of the stomach or mid-back?	691 / 3110	.247 ± .220 (–.280, .775)	.248 ± .220 (–.319, .816)	.155 ± .254 (–.515, .825)	.578 ± .456 (–.625, 1.78)	.42
Barium enema?	821 / 3149	.183 ± .224 (–.353, .718)	.182 ± .224 (–.395, .758)	.338 ± .252 (–.326, 1.00)	–.312 ± .493 (–1.61, .990)	.22
Upper GI?	1140 / 3167	.204 ± .222 (–.327, .735)	.201 ± .222 (–.370, .773)	.256 ± .275 (–.470, .981)	.105 ± .374 (–.881, 1.09)	.75
Intravenous pyelogram?	396 / 3147	.213 ± .223 (–.321, .746)	.222 ± .223 (–.352, .795)	.143 ± .239 (–.489, .774)	.919 ± .673 (–.855, 2.69)	.29
Fluoroscopy of the upper body?	246 / 3151	.193 ± .223 (–.341, .727)	.196 ± .223 (–.378, .771)	.239 ± .228 (–.363, .840)	–.459 ± .945 (–2.95, 2.03)	.46
Other nuclear scan?	216 / 3152	.239 ± .220 (–.287, .765)	.248 ± .219 (–.318, .813)	.325 ± .220 (–.255, .905)	–2.19 ± 1.53 (–6.22, 1.84)	.063
Dental x-rays that did not usually include a lead shield over the neck area?	1644 / 3181	.211 ± .220 (–.317, .738)	.207 ± .220 (–.361, .775)	.308 ± .309 (–.508, 1.12)	.108 ± .316 (–.726, .942)	.65

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

Table IX.P-40. Confounding and Effect Modification by Occupational History: Diffuse Ultrasound-Detected Abnormalities

Have You Ever Worked in Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted Estimate	Incl. Confounding Estimate	Including Effect Modification		P
				Group 0	Group 1	
Any metal industry?	238 / 3181	.211 ± .220 (–.317, .738)	.214 ± .220 (–.354, .782)	.253 ± .223 (–.337, .842)	–.669 ± 1.14 (–3.67, 2.33)	.41
Any nuclear facility?	370 / 3181	.211 ± .220 (–.317, .738)	.195 ± .222 (–.377, .767)	.096 ± .252 (–.568, .760)	.606 ± .479 (–.658, 1.87)	.36
Any other industry or occupation where you may have been exposed to radioactive materials or x-rays?	442 / 3181	.211 ± .220 (–.317, .738)	.219 ± .220 (–.348, .786)	.327 ± .237 (–.299, .953)	–.436 ± .698 (–2.28, 1.41)	.25
Any of the above industries or occupations?	891 / 3181	.211 ± .220 (–.317, .738)	.200 ± .221 (–.370, .770)	.283 ± .266 (–.417, .984)	.023 ± .403 (–1.04, 1.09)	.59

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

Table IX.P-41 displays the results of analyses of possible confounding or effect modification by smoking variables. There was no evidence that the dose-response was significantly confounded by either smoking variable, or that there was a dose-response that differed significantly according to smoking history.

Table IX.P-41. Confounding and Effect Modification by Smoking: Diffuse Ultrasound-Detected Abnormalities

Have You Ever Smoked Any of the Following: (0=No, 1=Yes)	Yes / Total	Estimated Dose-Response Coefficient (per Gy)				
		Unadjusted Estimate	Incl. Confounding. Estimate	Including Effect Modification		P
				Group 0	Group 1	
Cigarettes (unfiltered or filtered)?	1850 / 3173	.218 ± .220 (–.309, .744)	.224 ± .221 (–.344, .792)	.285 ± .366 (–.680, 1.25)	.190 ± .278 (–.545, .924)	.84
Any of cigarettes, cigar or pipe?	1896 / 3173	.218 ± .220 (–.309, .744)	.224 ± .221 (–.345, .793)	.217 ± .370 (–.759, 1.19)	.228 ± .275 (–.497, .953)	.98

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified logistic regression models. P-values indicate the statistical significance of two-tailed comparison of estimated coefficients between Groups 0 and 1.

P.5.i. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for the outcome of diffuse thyroid UDA are shown in Figure IX.P-8 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in that figure are calculated at the 98.33% confidence level, i.e., are adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates. While the point estimate of the slope is greater than 0 for 88 of the 100 realizations, the confidence interval includes 0 for all but 1 of the realizations. Also shown in Figure IX.P-8 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean and mean of each participant's 100 dose realizations. In summary, for only one of the 100 realizations of the estimated doses was there a statistically significant dose-response, although for most of the realizations the estimated slope was greater than 0.

Figure IX.P-8. Plot of Estimated Slope and 95% CI by Dose Realization: Diffuse Ultrasound-Detected Abnormalities

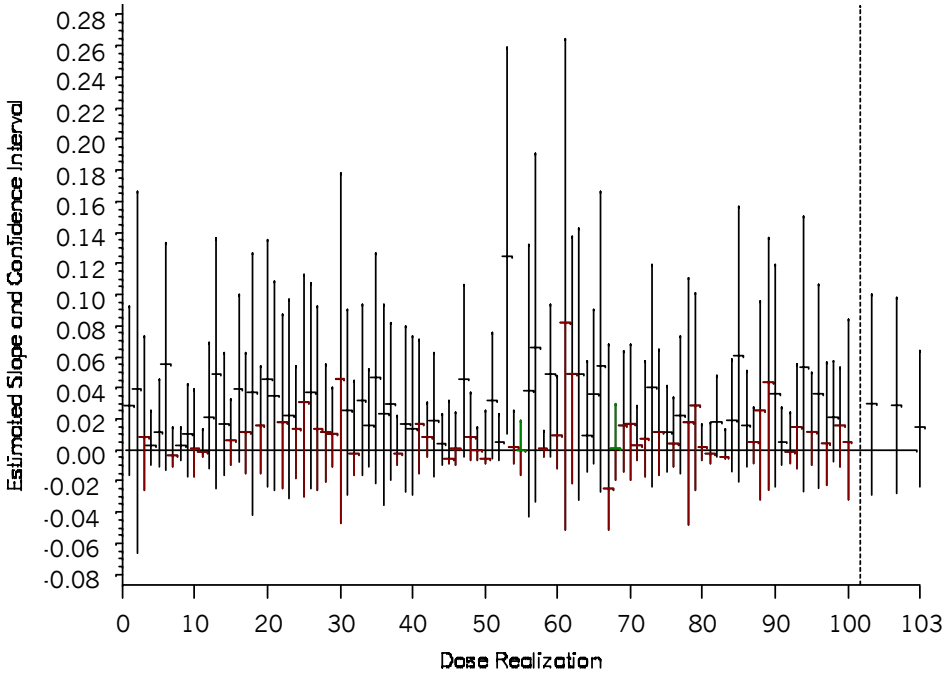
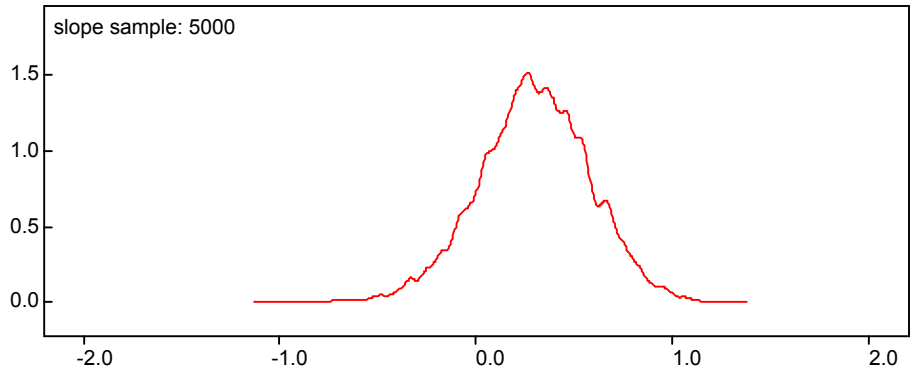


Figure IX.P-9 displays the distribution of the 5000 logistic regression coefficient estimates obtained by the simulation procedures described in section VIII.C.2.b.3 above. It is evident from the figure that most of the estimates were between about –0.5 and 1.0. The estimate was less than or equal to 0 for 713 of the 5000 replications, implying an empirical one-tailed p-value of 0.14. The median estimate was 0.30, and the upper and lower percentiles corresponding to the Bonferroni-adjusted 95% confidence interval were –0.42 and 0.96. These may be compared to the p-value of 0.17 and the estimate of 0.21 with confidence interval (–0.32, 0.74) obtained using the median dose estimates without adjustment for uncertainty. Thus, this method of adjusting the estimated logistic regression coefficient for the uncertainty in the dose estimates did not provide evidence that the prevalence of diffuse thyroid UDA increased with increasing dose.

Figure IX.P-9. Distribution of Simulation Estimates of Logistic Regression Coefficient: Diffuse Ultrasound-Detected Abnormalities



Q. Laboratory Values

Associations between laboratory values and estimated thyroid radiation dose were investigated by fitting the linear dose-response model [4], described in section VIII.C.1.a above. The regression coefficient B in this model plays a role analogous to that in the model [1] for cumulative incidence. In particular the direction and magnitude of each estimated dose-response relationship is represented by the estimate of the regression coefficient. An estimate of B greater than 0 indicates that the mean of the laboratory value tended to increase with increasing dose, while an estimate less than 0 indicates that the mean tended to decrease with increasing dose. The statistical significance of the dose-response was tested using the likelihood ratio statistic.

The p-values used to characterize the statistical significance of associations between lab values and estimated radiation dose were reported for two-tailed tests. This differed from the use of one-sided p-values in the tests for association with disease outcomes.

Of the 3191 living evaluable in-area participants, 3183 (99.7%) consented to provide a blood specimen at their HTDS clinic.

Q.1. Thyroid Stimulating Hormone (TSH)

Of the 3183 living evaluable in-area participants who provided blood samples, 222 were receiving exogenous thyroid hormone at the time of their HTDS clinic. These 222 were excluded from the analyses of TSH levels. Among the remaining 2961 living evaluable participants, 584 had TSH measured by RIA, 810 by EIA-1, and 1567 by EIA-2. Table IX.Q-1 displays the minimum, maximum, and median TSH levels of the 584 participants for whom RIA was used.

Table IX.Q-1. Distributions of TSH Levels Measured by RIA, by Sex

	TSH (μIU/ml) measured by RIA		
	Female (N = 281)	Male (N = 303)	Total (N = 584)
Minimum	0.1	0.3	0.1
Maximum	52.9	100.0	100.0
Median	2.3	2.3	2.3

Tables IX.Q-2 and IX.Q-3 display similar results for the participants whose TSH levels were measured by either of the two EIA assays. For two participants with TSH measured by EIA-1 and six with TSH measured by EIA-2, the TSH levels were reported simply as < 0.03 μIU/ml and < 0.04 μIU/ml, respectively. Such measurements are “left-censored”, that is, their specific values are not known, and they are known only to be less than the specified value.

Table IX.Q-2. Distributions of TSH Levels Measured by EIA-1, by Sex

	TSH (μIU/ml) measured by EIA-1		
	Female (N = 376)	Male (N = 434)	Total (N = 810)
Minimum	< 0.03	0.21	< 0.03
Maximum	50.34	28.77	50.34
Median	1.59	1.37	1.49
Number (%) below lower measurement limit (< 0.03 μIU/ml)	2 (0.5%)	0 (0%)	2 (0.2%)

Table IX.Q-3. Distributions of TSH Levels Measured by EIA-2, by Sex

	TSH (μIU/ml) measured by EIA-2		
	Female (N = 766)	Male (N = 801)	Total (N = 1567)
Minimum	< 0.04	< 0.04	< 0.04
Maximum	24.12	22.46	24.12
Median	1.49	1.22	1.35
Number (%) below lower measurement limit (< 0.04 μIU/ml)	5 (0.7%)	1 (0.1%)	6 (0.4%)

It is evident from Tables IX.Q-1 through IX.Q-3 above that the distributions of TSH values were quite skewed to the right, since the median values are much closer to the minima than the maxima. Therefore the regression model was applied to the logarithms of the TSH values. Figures IX.Q-1 through IX.Q.3 display the TSH values, plotted on the logarithmic scale, in relation to estimated thyroid radiation dose.

Figure IX.Q-1. Scatter Plot of TSH by RIA and Estimated Dose

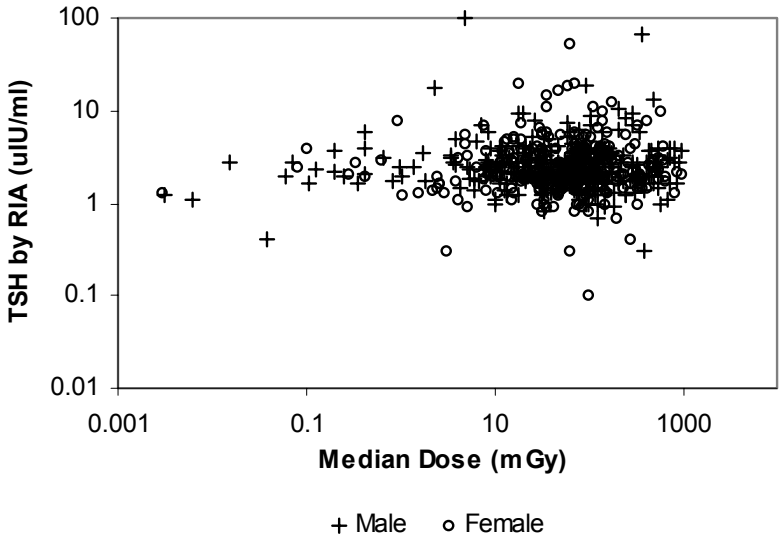


Figure IX.Q-2. Scatter Plot of TSH by EIA-1 and Estimated Dose

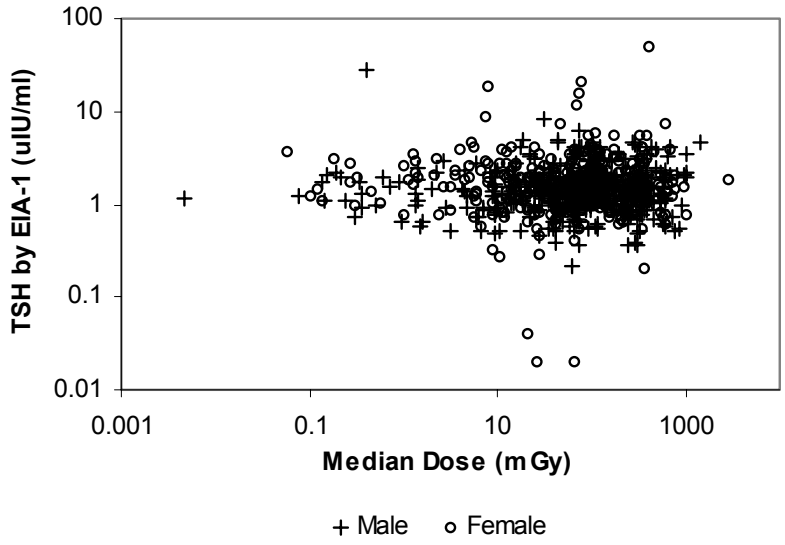
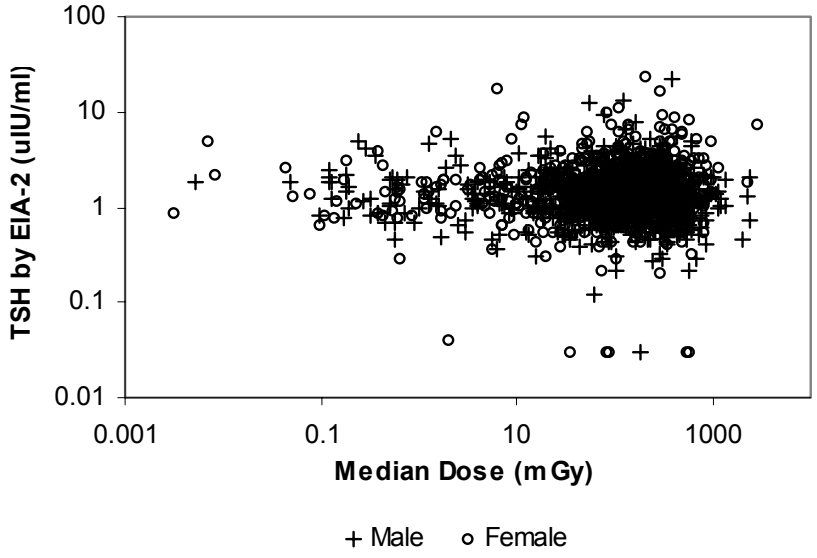


Figure IX.Q-3. Scatter Plot of TSH by EIA-2 and Estimated Dose



The results of fitting the sex-stratified linear dose-response model for log(TSH) are summarized for the three types of assays in Table IX.Q-4 below. In the table, parameter estimates are converted from the scale of log(TSH) back to TSH. So, for example, the estimated average RIA-based TSH for women of 2.45 μIU/ml is in fact an estimate of the geometric mean. Also, the radiation dose effect is represented by the percentage change per Gy. For example, the estimated average (geometric mean) TSH level for women based on EIA-1 increased from 1.58 μIU/ml at 0 Gy to $1.58 \times 1.142^2 = 2.06$ μIU/ml at 2 Gy (2000 mGy). For none of the three assays was there a significant trend in relation to estimated radiation dose.

Table IX.Q-4. Parameter Estimates for Dose-Response Models: TSH

Parameter	RIA (μIU/ml)	EIA – 1 (μIU/ml)	EIA – 2 (μIU/ml)
No. of living evaluable participants	584	810	1567
No. with left-censored values	0	2	6
Estimated average background TSH for women	2.45 (2.22, 2.71)	1.58 (1.45, 1.72)	1.48 (1.39, 1.57)
Estimated average background TSH for men	2.43 (2.20, 2.68)	1.36 (1.26, 1.48)	1.26 (1.18, 1.34)
Estimated percentage change in average TSH per Gy	+2.0% (–30.0%, +48.9%)	+14.2% (–12.0%, +48.0%)	+1.5% (–12.2%, +17.2%)
Statistical significance of dose-response (two-tailed p-value)	0.90	0.22	0.82

Entries in the table for model parameters are the parameter estimate, with Bonferroni-adjusted 95% confidence interval in parentheses, based on sex-stratified linear model for log (TSH).

Since the average levels of TSH differed rather substantially among the three assays, it was not considered appropriate to simply combine all three groups and attempt to fit the simple sex-stratified linear regression model [4]. Therefore a generalization of the sex-stratified linear model was examined, in which the mean values of log(TSH) were assumed to differ between the sexes and according to the type of assay. When this model was fit to the data for all 2961 living evaluable participants with TSH measurements, there was still no significant trend of average log(TSH) in relation to estimated thyroid radiation dose. If a common slope was assumed for all three assays, the estimated regression coefficient was +4.5% per Gy with Bonferroni-adjusted 95% confidence limits –8.4% and +19.1% per Gy, which was not significantly different from zero (two-tailed p = 0.42).

Q.1.a. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for log (TSH) by RIA, EIA-1, and EIA-2 are shown in Figures IX.Q-4 through IX.Q-6 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in these figures were calculated at the 98.33% confidence level, i.e., were adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates. The point estimate of the slope was greater than 0 for 59, 94, and 65 of the 100 realizations for TSH by RIA, EIA-1, and EIA-2, respectively. However the confidence intervals included 0 for all 100 realizations for TSH by RIA and EIA-2, and for 96 of the 100 realizations for EIA-1.

Also shown in Figures IX.Q-4 through IX.Q-6 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean, and mean of each participant's 100 dose realizations.

Figure IX.Q-4. Estimated Dose-Response for TSH by RIA, by Dose Realization

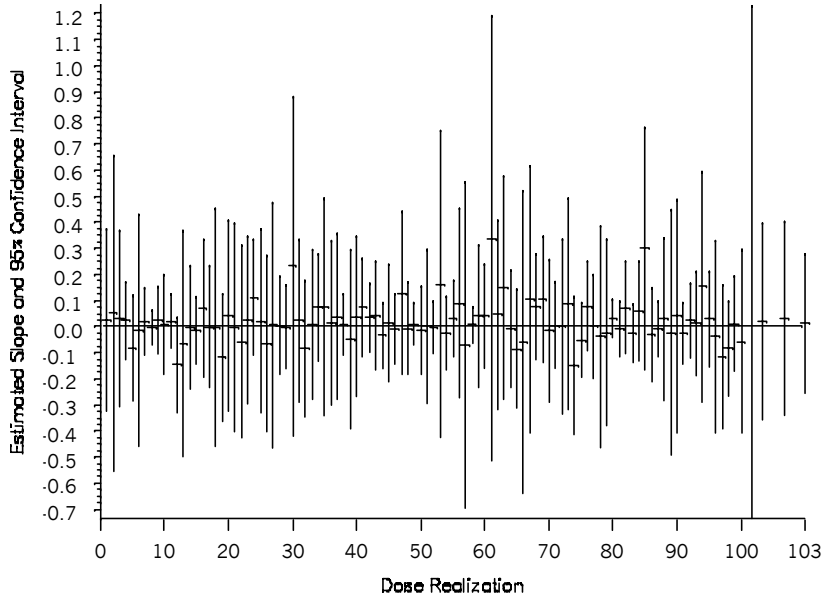


Figure IX.Q-5. Estimated Dose-Response for TSH by EIA-1, by Dose Realization

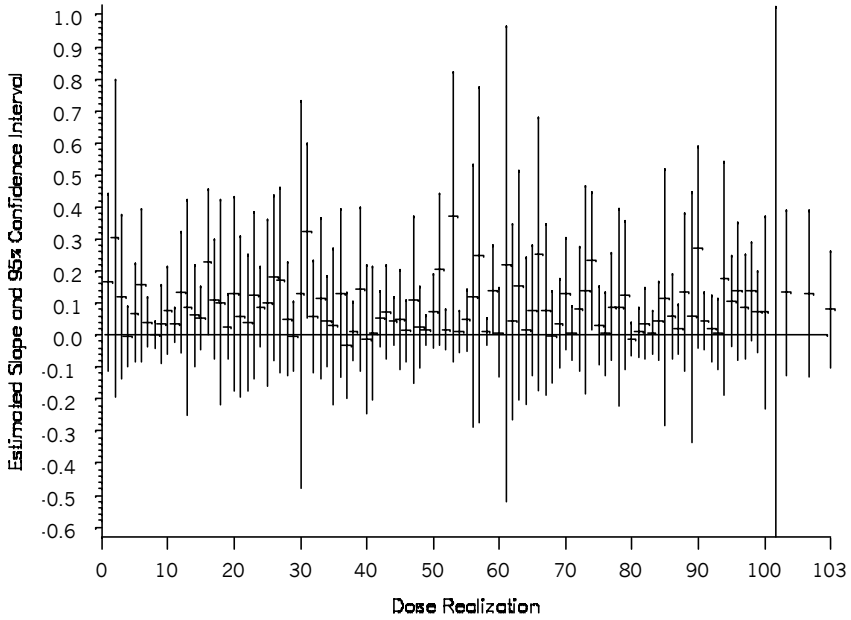
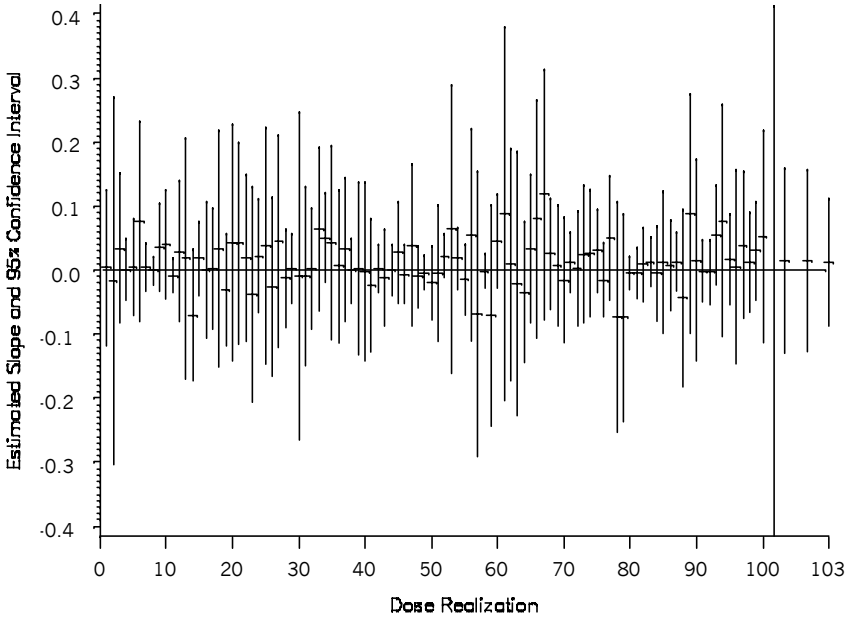


Figure IX.Q-6. Estimated Dose-Response for TSH by EIA-2, by Dose Realization



Q.2. Total Thyroxine (T4)

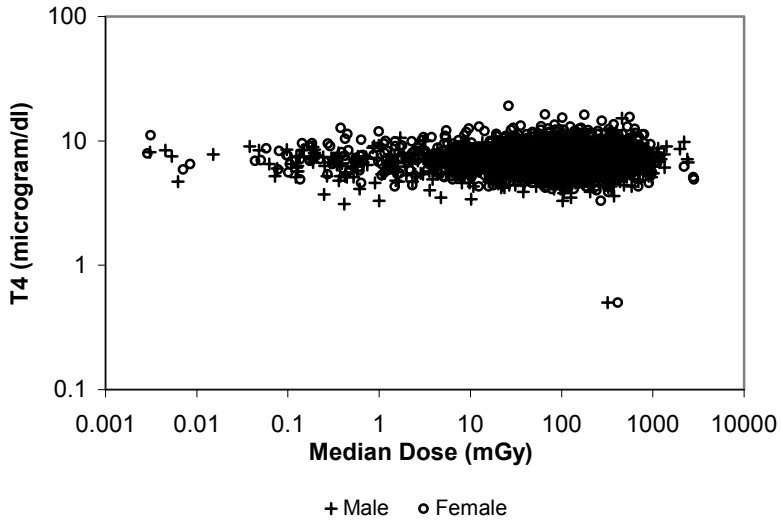
The 222 living evaluable in-area participants who were receiving exogenous thyroid hormone at the time of their HTDS clinic were excluded from the primary analysis of T4. The T4 values were unknown for two additional in-area participants due to insufficient volumes of collected blood. Table IX.Q-5 displays the minimum, maximum, and median T4 levels of the 2959 participants for whom data were available. For two of these participants the T4 levels were left censored, reported as < 1.0 µg/dl. All other T4 levels were 3.1 µg/dl or greater. Therefore the distribution of T4 levels was somewhat skewed to the right, and consequently the regression model was applied to the logarithms of the T4 values.

Table IX.Q-5. Distributions of Total Thyroxine (T4) Levels, by Sex

	T4 (µg/dl)		
	Female (N = 1422)	Male (N = 1537)	Total (N = 2959)
Minimum	< 1.0	< 1.0	< 1.0
Maximum	19.1	15.2	19.1
Median	7.5	6.6	7.0
Number (%) below lower measurement limit (< 1.0 µg/dl)	1 (0.07%)	1 (0.07%)	2 (0.07%)

T4 values, plotted on the logarithmic scale, are shown by estimated dose in Figure IX.Q-7.

Figure IX.Q-7. Scatter Plot of T4 and Estimated Dose



The results of fitting the sex-stratified linear dose-response model for log(T4) are summarized in Table IX.Q-6 below. In the table, parameter estimates are converted from the scale of log(T4) back to T4. So, for example, the estimated average T4 of 7.52 μg/dl for women is in fact an estimate of the geometric mean. Also, the radiation dose effect is represented by the percentage change per Gy. For example, the estimated average (geometric mean) T4 level for women decreased from 7.52 μg/dl at 0 Gy to $7.52 \times 0.996^2 = 7.46$ μg/dl at 2 Gy (2000 mGy). There was no significant trend of T4 in relation to estimated radiation dose (two-tailed p = 0.84).

Table IX.Q-6. Parameter Estimates for Dose-Response Models: T4

Parameter	T4 (μg/dl)
No. of living evaluable participants	2959
No. with left-censored values	2
Estimated average background T4 for women	7.52 (7.41, 7.64)
Estimated average background T4 for men	6.58 (6.48, 6.67)
Estimated percentage change in average T4 per Gy	–0.4% (–4.5%, +4.0%)
Statistical significance of dose-response (two-tailed p-value)	0.84

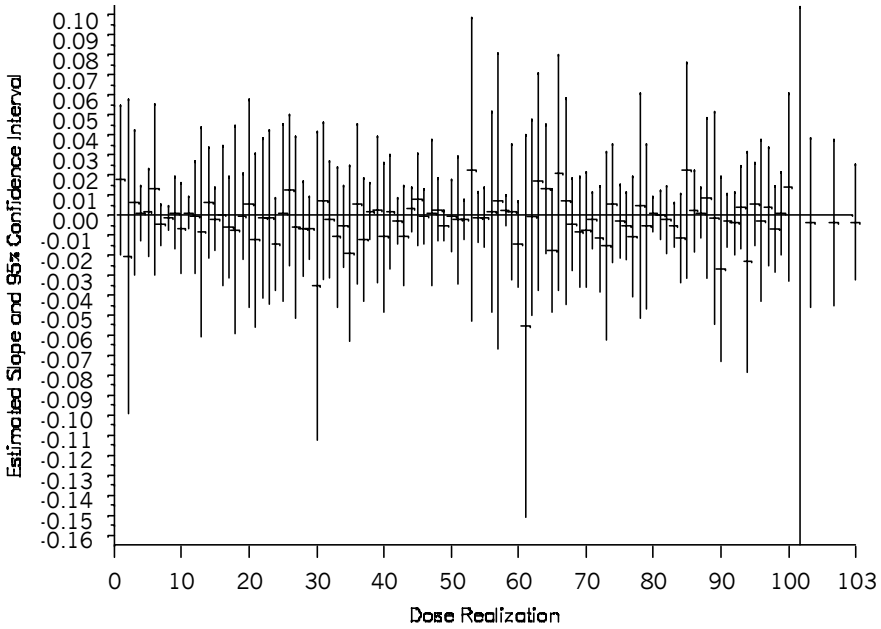
Entries for model parameters (background means and slope) in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses.

Q.2.a. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for log of T4 are shown in

Figure IX.Q-8 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in that figure were calculated at the 98.33% confidence level, i.e., were adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates. While the point estimate of the slope was greater than 0 for 42 of the 100 realizations, the confidence interval included 0 for all of the 100 realizations. Also shown in Figure IX.Q-8 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean, and mean of each participant's 100 dose realizations.

Figure IX.Q-8. Estimated Dose-Response for T4, by Dose Realization



Q.3. Triiodothyronine Resin Uptake (T3RU)

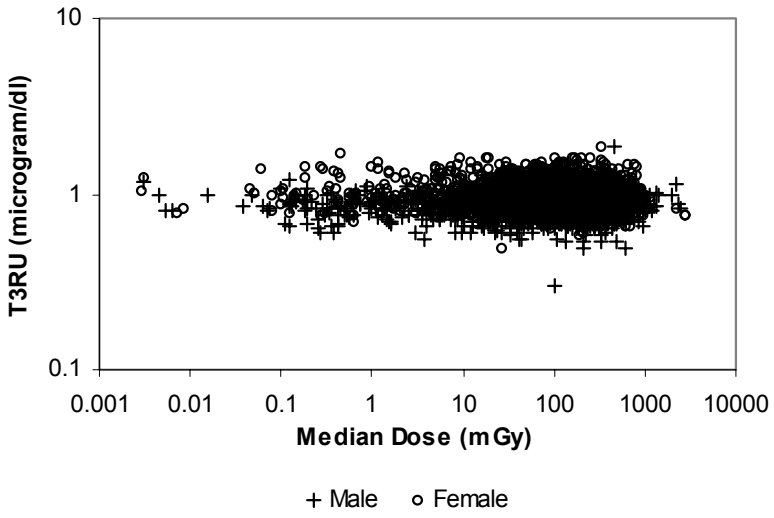
The 222 living evaluable in-area participants who were receiving exogenous thyroid hormone at the time of their HTDS clinic were excluded from the analyses of T3RU. The T3RU values were unknown for two additional in-area participants, the same two whose T4 values were unknown due to insufficient volumes of collected blood. Table IX.Q-7 displays the minimum, maximum, and median T3RU levels of the 2959 participants for whom data were available. For one of these participants the T3RU level was left censored, reported as < 0.4 µg/dl. All other T3RU levels were 0.49 µg/dl or greater. The distribution of T3RU levels was somewhat skewed to the right, and therefore the regression model was applied to the logarithms of the T3RU values.

Table IX.Q-7. Distributions of T3 Resin Uptake (T3RU), by Sex

	T3RU (μg/dl) Female (N = 1422)	Male (N = 1537)	Total (N = 2959)
Minimum	0.49	< 0.4	< 0.4
Maximum	1.87	1.86	1.87
Median	1.00	0.86	0.92
Number (%) below lower measurement limit (< 0.4)	0 (0%)	1 (0.1%)	1 (0.03%)

T3RU values, plotted on the logarithmic scale, are shown by estimated dose in Figure IX.Q-9.

Figure IX.Q-9. Scatter Plot of T3RU and Estimated Dose



The results of fitting the sex-stratified linear dose-response model for log(T3RU) are summarized in Table IX.Q-8 below. In the table, parameter estimates are converted from the scale of log(T3RU) back to T3RU. So, for example, the estimated average T3RU of 1.02 μg/dl for women is in fact an estimate of the geometric mean. Also, the radiation dose effect is represented by the percentage change per Gy. For example, the estimated average (geometric mean) T3RU level for women decreased from 1.02 μg/dl at 0 Gy to $1.02 \times 0.988^2 = 1.00$ μg/dl at 2 Gy (2000 mGy). There was no significant trend of T3RU in relation to estimated radiation dose (two-tailed p = 0.36).

Table IX.Q-8. Parameter Estimates for Dose-Response Models: T3RU

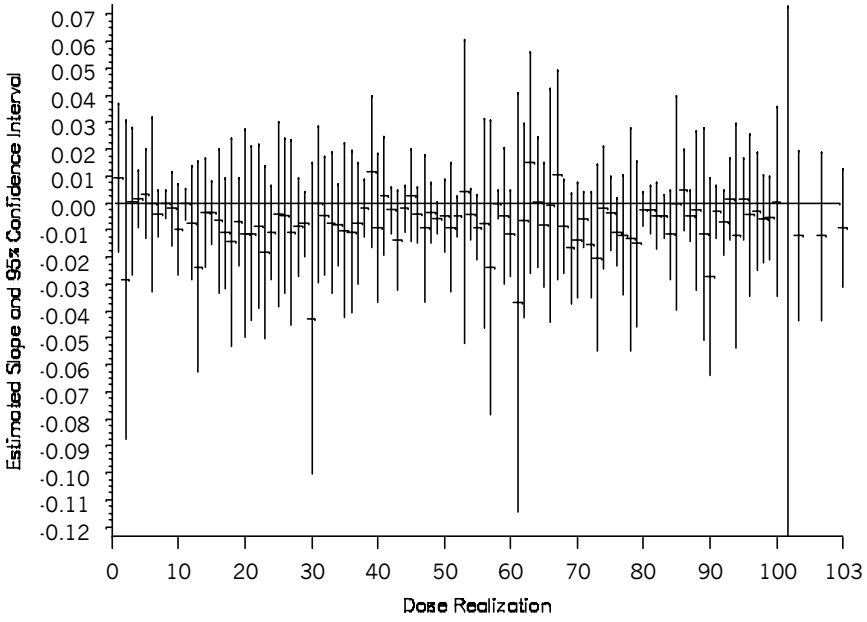
Parameter	T3RU (μg/dl)
No. of living evaluable participants	2959
No. with left-censored values	1
Estimated average background T3RU for women	1.02 (1.01, 1.03)
Estimated average background T3RU for men	0.85 (0.84, 0.86)
Estimated percentage change in average T3RU per Gy	–1.2% (–4.3%, +2.0%)
Statistical significance of dose-response (two-tailed p-value)	0.36

Entries for model parameters (background means and slope) in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses.

Q.3.a. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for log of T3RU values are shown in Figure IX.Q-10 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in that figure were calculated at the 98.33% confidence level, i.e., were adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates. While the point estimate of the slope was greater than 0 for 15 of the 100 realizations, the confidence interval included 0 for all of the 100 realizations. Also shown in Figure IX.Q-10 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean, and mean of each participant's 100 dose realizations.

Figure IX.Q-10. Estimated Dose-Response for T3RU, by Dose Realization



Q.4. Free Thyroxine Index (FTI)

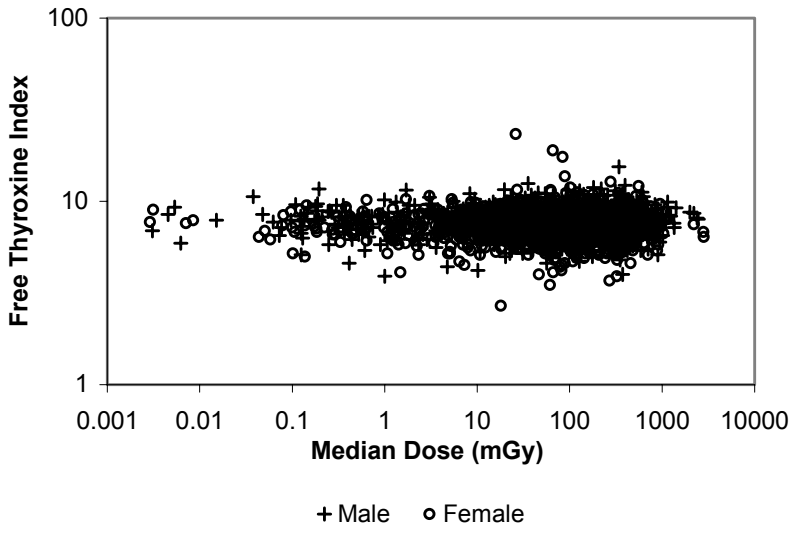
The 222 living evaluable in-area participants who were receiving exogenous thyroid hormone at the time of their HTDS clinic were excluded from the primary analyses of FTI. The FTI values were unknown for five additional in-area participants: the two with unknown T4 and T3RU, and three others for whom either T4 or T3RU was below its level of detection. Table IX.Q-9 displays the minimum, maximum, and median FTI values of the 2956 participants for whom data were available. Since the distribution of FTI values was somewhat skewed to the right, regression modeling of the dose-response was applied to the logarithms of the FTI values.

Table IX.Q-9. Distributions of Free Thyroxine Index (FTI), by Sex

	FTI		
	Female (N = 1421)	Male (N = 1535)	Total (N = 2956)
Minimum	2.7	3.9	2.7
Maximum	23.3	15.4	23.3
Median	7.4	7.8	7.6

FTI values, plotted on the logarithmic scale, are shown by estimated dose in Figure IX.Q-11.

Figure IX.Q-11. Scatter Plot of FTI and Estimated Dose



The results of fitting the sex-stratified linear dose-response model for log(FTI) are summarized in Table IX.Q-10 below. In the table, parameter estimates are converted from the scale of log(FTI) back to FTI. So, for example, the estimated average FTI of 7.38 for women is in fact an estimate of the geometric mean. Also, the radiation dose effect is represented by the percentage change per Gy. For example, the estimated average (geometric mean) FTI level for women increased from 7.38 at 0 Gy to $7.38 \times 1.016^2 = 7.62$ at 2 Gy (2000 mGy). There was no significant trend of FTI in relation to estimated radiation dose (two-tailed p = 0.23).

Table IX.Q-10. Parameter Estimates for Dose-Response Models: FTI

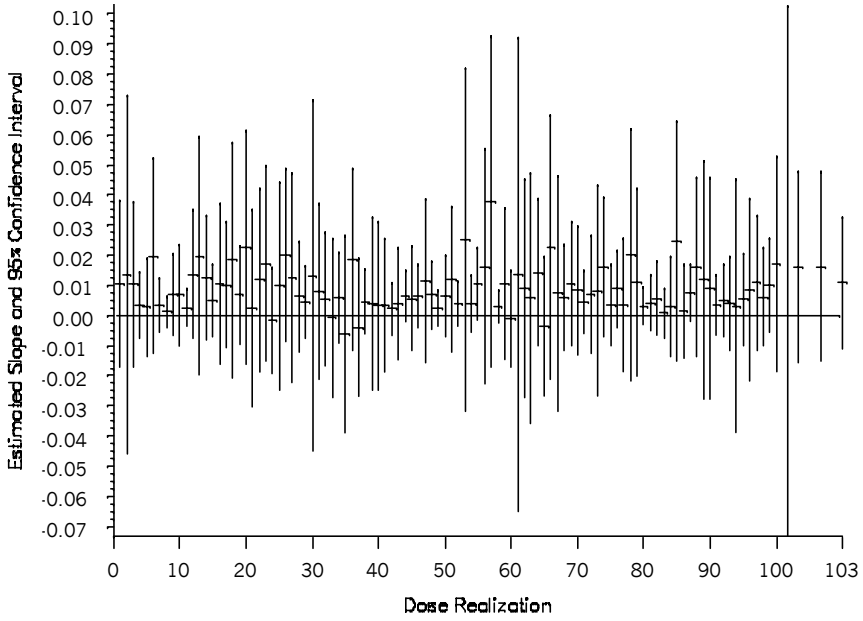
Parameter	FTI
No. of living evaluable participants	2956
Estimated average background FTI for women	7.38 (7.29, 7.46)
Estimated average background FTI for men	7.72 (7.63, 7.81)
Estimated percentage change in average FTI per Gy	+1.6% (−1.6%, +4.9%)
Statistical significance of dose-response (two-tailed p-value)	0.23

Entries for model parameters (background means and slope) in the tableare estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses.

Q.4.a. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for log of FTI are shown in Figure IX.Q-12 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in that figure were calculated at the 98.33% confidence level, i.e., were adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates. While the point estimate of the slope was greater than 0 for 94 of the 100 realizations, the confidence interval included 0 for all of the 100 realizations. Also shown in Figure IX.Q-12 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean, and mean of each participant's 100 dose realizations.

Figure IX.Q-12. Estimated Dose-Response for FTI, by Dose Realization



Q.5. Anti-Thyroid Autoimmune Response

Anti-TPO or AMA values were used to measure the anti-thyroid autoimmune responses of 1562 and 1620 in-area living evaluable participants, respectively. Neither assay result was available for eight participants who declined to provide a blood sample, and for one other whose sample was of insufficient volume. Tables IX.Q-11 and IX.Q-12 display the minimum, maximum, and median anti-TPO or AMA values of the participants for whom data were available. For both assays, the majority of participants had values below the lower measurement limits: 80% with anti-TPO < 2.0 IU/ml, and 78% with AMA < 20 IU/ml. In addition, 6% of the participants assayed by AMA had values above the upper measurement limit, i.e., > 700 U/ml. Since the distributions of these values were skewed to the right, they were log-transformed for regression modeling of the dose-responses.

Table IX.Q-11. Distributions of Anti-TPO, by Sex

	Anti-TPO (IU/ml)		
	Female (N = 812)	Male (N = 750)	Total (N = 1562)
Minimum	< 2.0	< 2.0	< 2.0
Maximum	9569.7	1631.7	9569.7
Median	< 2.0	< 2.0	< 2.0
Number (%) below lower measurement limit (< 2.0 IU/ml)	594 (73%)	651 (87%)	1245 (80%)

Table IX.Q-12. Distributions of AMA, by Sex

	AMA (U/ml)		
	Female (N = 803)	Male (N = 817)	Total (N = 1620)
Minimum	< 20	< 20	< 20
Maximum	> 700	> 700	> 700
Median	< 20	< 20	< 20
Number (%) below lower measurement limit (< 20 U/ml)	590 (73%)	674 (82%)	1264 (78%)
Number (%) above upper measurement limit (> 700 U/ml)	63 (8%)	30 (4%)	93 (6%)

Anti-TPO and AMA results, plotted on logarithmic scales, are shown by estimated dose in Figures IX.Q-13 and IX.Q-14.

Figure IX.Q-13. Scatter Plot of Anti-TPO and Estimated Dose

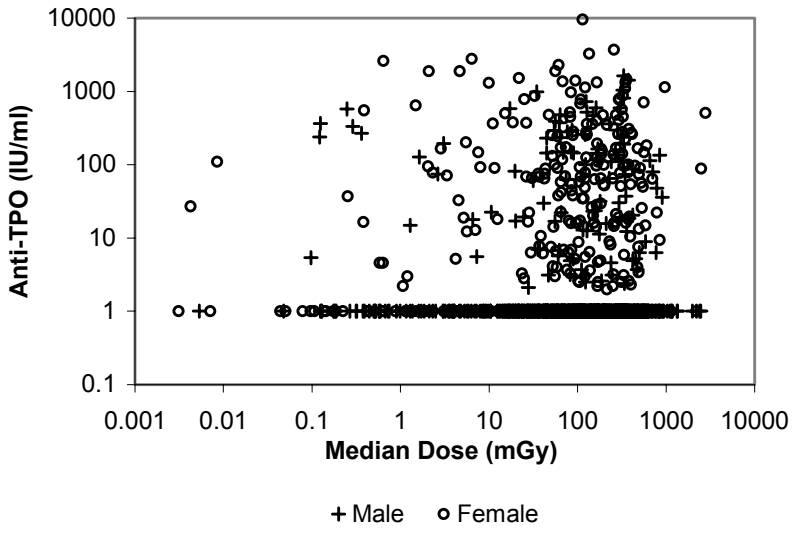
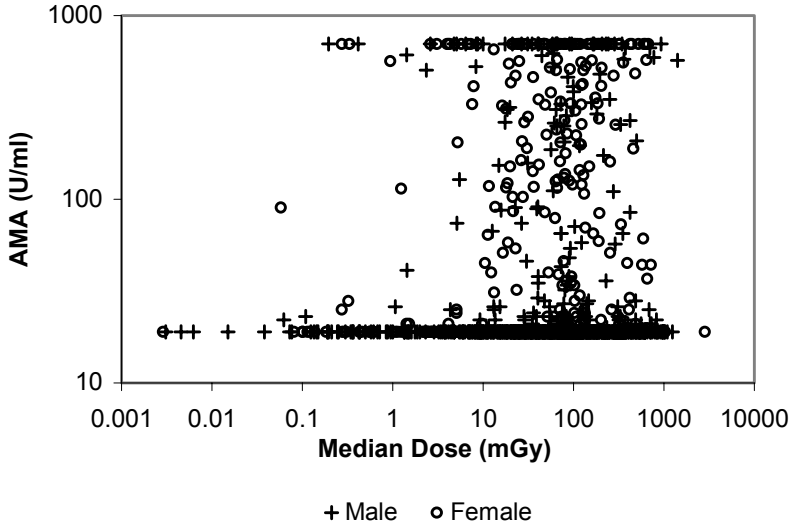


Figure IX.Q-14. Scatter Plot of AMA and Estimated Dose



The results of fitting the sex-stratified linear dose-response models for log(anti-TPO) and log(AMA) are summarized in Table IX.Q-13 below. In the table, parameter estimates are converted from the scales of logarithmically transformed values back to the original scales. So, for example, the estimated average anti-TPO of 0.03 IU/ml for women is in fact an estimate of the geometric mean. Also, the radiation dose effect is represented by the percentage change per Gy, with estimates less than zero indicating that the assay value decreased with increasing estimated thyroid dose. Since the majority of participants had anti-TPO or AMA values below their respective lower limits of measurement, the estimated parameter values have little meaning. Nevertheless these regression results provide no evidence that either value tended to increase sharply with increasing estimated dose (two-tailed p = 0.66 for anti-TPO, p = 0.52 for AMA).

Table IX.Q-13. Parameter Estimates for Dose-Response Models: Anti-TPO and AMA

Parameter	Anti-TPO (IU/ml)	AMA (U/ml)
No. of living evaluable participants	1562	1620
No. with left-censored values	1245	1264
No. with right-censored values	0	93
Estimated average background for women	0.03 (0.01, 0.07)	1.30 (0.63, 2.68)
Estimated average background for men	0.001 (0.000, 0.005)	0.29 (0.12, 0.70)
Estimated percentage change in average per Gy	–32.1% (–91.7%, +453%)	–39.9% (–91.2%, +312%)
Statistical significance of dose-response (two-tailed p-value)	0.66	0.52

Entries for model parameters (background means and slope) in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses.

Q.5.a. Uncertainty

The estimated slopes of the sex-stratified linear dose-response models for log of Anti-TPO and log of AMA are shown in Figures IX.Q-15 and IX.Q-16 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in these figures were calculated at the 98.33% confidence level, i.e., were adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates. While the point estimate of the slope was greater than 0 for 30 of the 100 realizations for anti-TPO and for 18 realizations for AMA, the confidence intervals for both assays included 0 for all of the 100 dose realizations. Also shown in Figures IX.Q-15 and IX.Q-16 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean, and mean of each participant's 100 dose realizations.

Figure IX.Q-15. Estimated Dose-Response for Anti-TPO, by Dose Realization

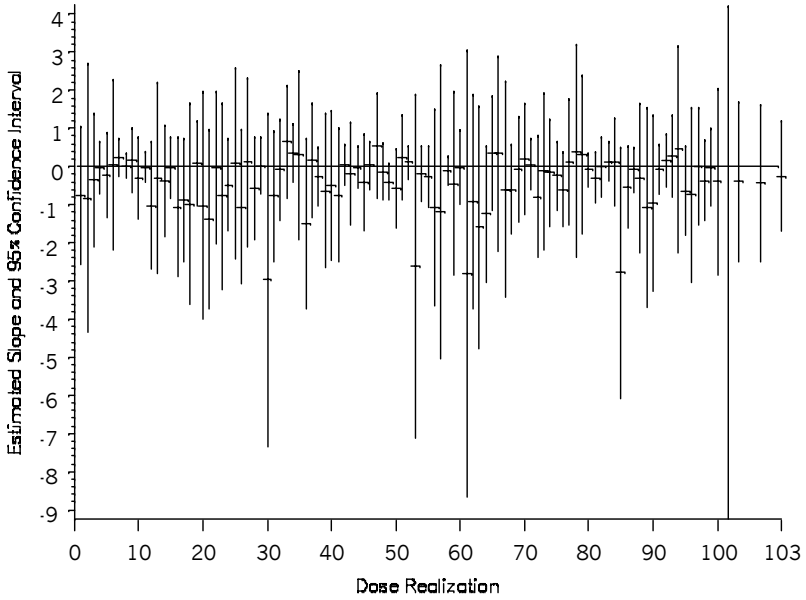
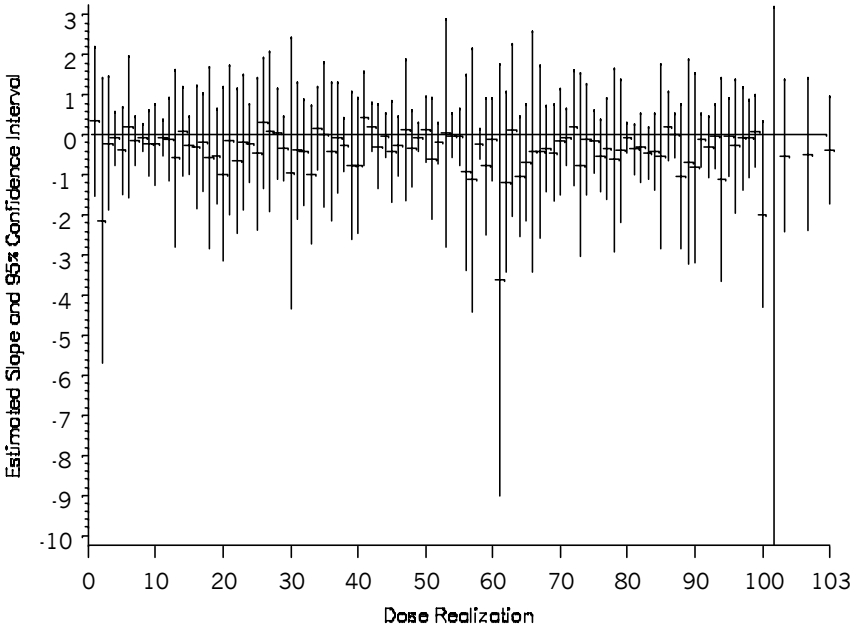


Figure IX.Q-16.Estimated Dose-Response for AMA, by Dose Realization



Q.6. Anti-Thyroglobulin Antibody (anti-TG)

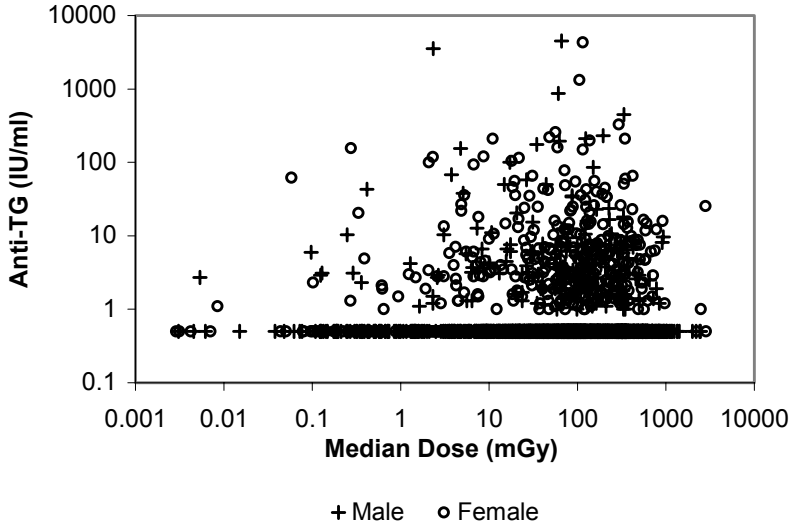
Anti-TG values were not available for 13 of the 3183 living evaluable in area participants who provided blood specimens due to insufficient volume and broken tubes. Table IX.Q-14 displays the minimum, maximum, and median anti-TG values of the 3170 participants for whom data were available. The lower limit of measurement for anti-TG was 1.0 IU/ml, and the majority of participants (85%) had values reported as < 1.0 IU/ml. Since the distribution of anti-TG values was skewed to the right, regression modeling of the dose-response was applied to the logarithms of the anti-TG values.

Table IX.Q-14. Distributions of Anti-TG, by Sex

	Anti-TG (IU/ml)		
	Female (N = 1607)	Male (N = 1563)	Total (N = 3170)
Minimum	< 1.0	< 1.0	< 1.0
Maximum	4300	4500	4500
Median	< 1.0	< 1.0	< 1.0
Number (%) below lower measurement limit (< 1.0 IU/ml)	1281 (80%)	1400 (90%)	2681 (85%)

Anti-TG values, plotted on a logarithmic scale, are shown by estimated dose in Figure IX.Q-17. For clarity, the 2681 values that were below the lower measurement limit of 1.0 IU/ml are plotted at 0.5 IU/ml.

Figure IX.Q-17. Scatter Plot of Anti-TG and Estimated Dose



The results of fitting the sex-stratified linear dose-response model for log(anti-TG) are summarized in Table IX.Q-15 below. In the table, parameter estimates are converted from the scale of log(anti-TG) back to anti-TG. So, for example, the estimated average anti-TG of 0.02 IU/ml for women is in fact an estimate of the geometric mean. Also, the radiation dose effect is represented by the percentage

change per Gy, with the estimate less than zero indicating that the average anti-TG level decreased with increasing estimated thyroid dose. Since the majority of participants had anti-TG values below the lower limit of measurement, the estimated parameter values have little meaning. Nevertheless these regression results provide no evidence that average anti-TG levels tended to increase sharply with increasing estimated dose (two-tailed p = 0.20).

Table IX.Q-15. Parameter Estimates for Dose-Response Models: Anti-TG

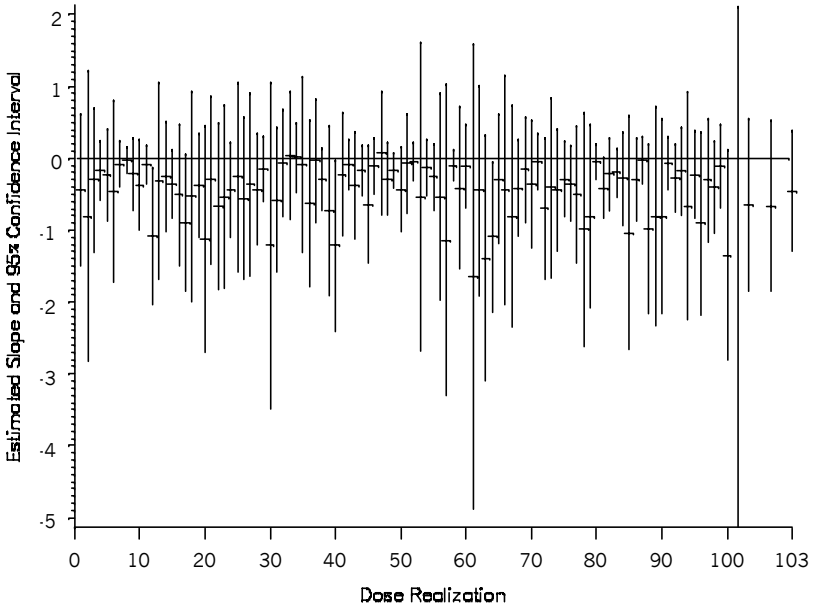
Parameter	Anti-TG (IU/ml)
No. of living evaluable participants	3170
No. with left-censored values	2681
Estimated average background anti-TG for women	0.02 (0.01, 0.03)
Estimated average background anti-TG for men	0.003 (0.002, 0.007)
Estimated percentage change in anti-TG Average per Gy	–47.3% (–84.2%, +75.5%)
Statistical significance of dose-response (two-tailed p-value)	0.20

Entries for model parameters (background means and slope) in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses.

Q.6.a. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for log of anti-TG are shown in Figure IX.Q-18 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in that figure were calculated at the 98.33% confidence level, i.e., were adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates. While the point estimate of the slope was greater than 0 for 3 of the 100 realizations, the confidence interval included 0 for 97 of the 100 realizations. Also shown in Figure IX.Q-18 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean, and mean of each participant's 100 dose realizations.

Figure IX.Q-18. Estimated Dose-Response for Anti-TG, by Dose Realization



Q.7. Serum Calcium

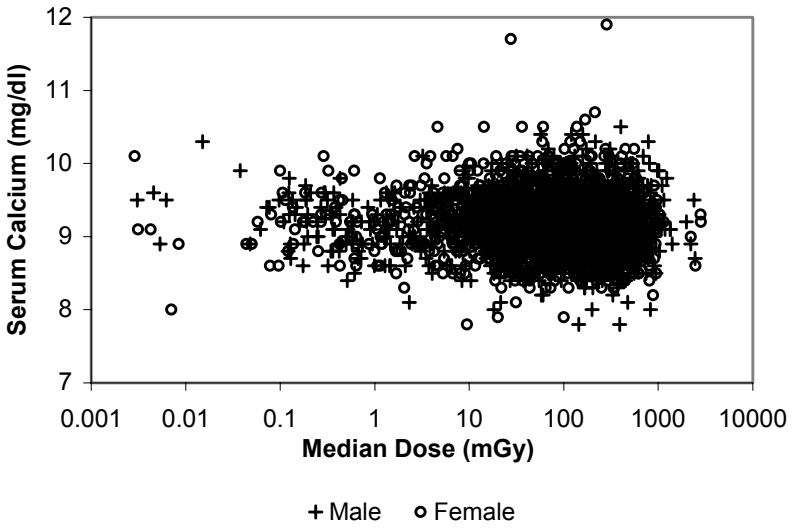
Of the 3183 living evaluable in-area participants who provided blood samples, 227 with diagnoses of hypothyroidism or hyperthyroidism based on the HTDS examination were excluded from the primary analysis of serum calcium levels. Two additional participants did not have serum calcium data due to insufficient volumes of collected blood. Table IX.Q-16 displays the minimum, maximum, and median serum calcium levels of the 2954 participants for whom data were available.

Table IX.Q-16. Distributions of Serum Calcium, by Sex

	Serum Calcium (mg/dl)		
	Female (N = 1448)	Male (N = 1506)	Total (N = 2954)
Minimum	7.8	7.8	7.8
Maximum	11.7	10.5	11.7
Median	9.1	9.2	9.2

Serum calcium levels are shown by estimated dose in Figure IX.Q-19.

Figure IX.Q-19. Scatter Plot of Serum Calcium and Estimated Dose



As can be seen in Figure IX.Q-19, the overall distribution of serum calcium levels was fairly symmetrically distributed, therefore the sex-stratified linear dose-response model [4] was fit without logarithmic transformation. The results are summarized in Table IX.Q-17 below. There was a statistically significant trend of decreasing serum calcium level in relation to increasing radiation dose ($p = 0.0074$). The estimated background means were 9.17 mg/dl for female and 9.19 mg/dl for male, with Bonferroni-adjusted 95% confidence intervals (9.14, 9.20) and (9.16, 9.22), respectively. The estimated slope of the dose-response was –0.09 mg/dl per Gy, with confidence interval ranging from –0.16 to –0.01 mg/dl per Gy, implying that the mean decreased by an average of 0.09 mg/dl with each incremental dose of 1 Gy (1000 mGy). Although this trend is statistically significant, it is small enough in magnitude that the average serum calcium levels remain within the normal range of 8.4 – 10.2 mg/dl. For example, at 3 Gy (3000 mGy), which is larger than the largest dose estimate of any study participant, the average serum calcium level predicted by the regression model for female is $9.17 - 0.09 \times 3 = 8.90$ mg/dl.

Table IX.Q-17. Parameter Estimates for Linear Dose-Response Models: Serum Calcium

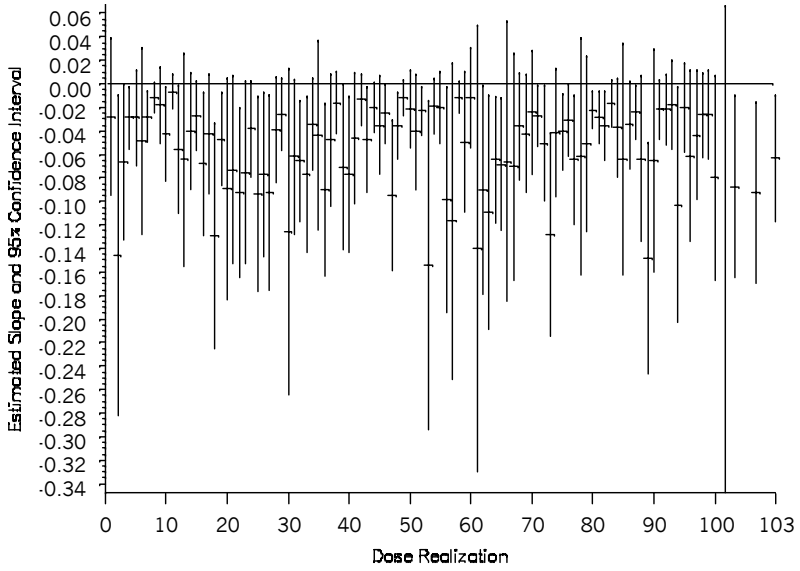
Parameter	Serum Calcium (mg/dl)
No. of living evaluable participants	2954
Estimated average background serum calcium for women	9.17 ± .01 (9.14, 9.20)
Estimated average background serum calcium for men	9.19 ± .01 (9.16, 9.22)
Estimated slope of dose-response (per Gy)	–.09 ± .03 (-.16, -.01)
Statistical significance of dose-response (two-tailed p-value)	0.0074

Entries for model parameters (background means and slope) in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses.

Q.7.a. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for serum calcium are shown in Figure IX.Q-20 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in that figure were calculated at the 98.33% confidence level, i.e., were adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates. While the point estimate of the slope was less than 0 for all 100 realizations, the confidence interval included 0 for 61 of the 100 realizations. Also shown in Figure IX.Q-20 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean, and mean of each participant's 100 dose realizations.

Figure IX.Q-20. Estimated Dose-Response for Serum Calcium, by Dose Realization



Q.8. Thyroid Mass

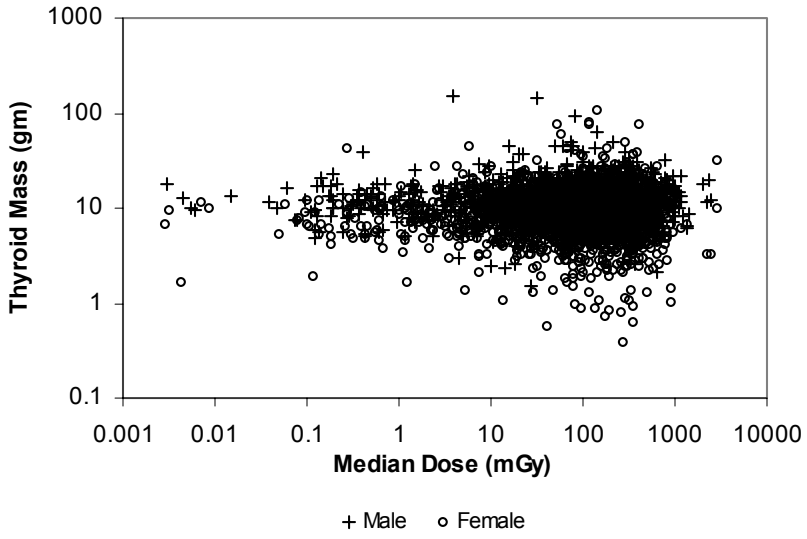
Estimates of thyroid mass were available for 3153 were in-area participants. Table IX.Q-18 displays the minimum, maximum, and median estimates of thyroid mass of these 3153 participants.

Table IX.Q-18. Distributions of Estimated Thyroid Mass, by Sex

	Thyroid Mass (gm)		
	Female (N = 1592)	Male (N = 1561)	Total (N = 3153)
Minimum	0.39	1.53	0.39
Maximum	108.62	149.78	149.78
Median	7.81	11.4	9.53

Thyroid mass, plotted on the logarithmic scale, is shown by estimated dose in Figure IX.Q-21 for these 3153 living evaluable in-area participants.

Figure IX.Q-21. Scatter Plot of Estimated Thyroid Mass and Estimated Dose



The results of fitting the sex-stratified linear dose-response model for log-transformed values of thyroid mass are summarized in Table IX.Q-19 below. In the table, parameter estimates are converted from the logarithmically transformed scale back to the scale of thyroid mass in grams. So, for example, the estimated average thyroid mass of 7.69 gm for women is in fact an estimate of the geometric mean. Also, the radiation dose effect is represented by the percentage change per Gy. For example, the estimated average (geometric mean) thyroid mass level for women decreased from 7.69 gm at 0 Gy to $7.69 \times 0.999^2 = 7.67$ gm at 2 Gy (2000 mGy). There was no significant trend of thyroid mass in relation to estimated radiation dose (two-tailed $p = 0.98$).

Table IX.Q-19. Parameter Estimates for Dose-Response Models: Thyroid Mass

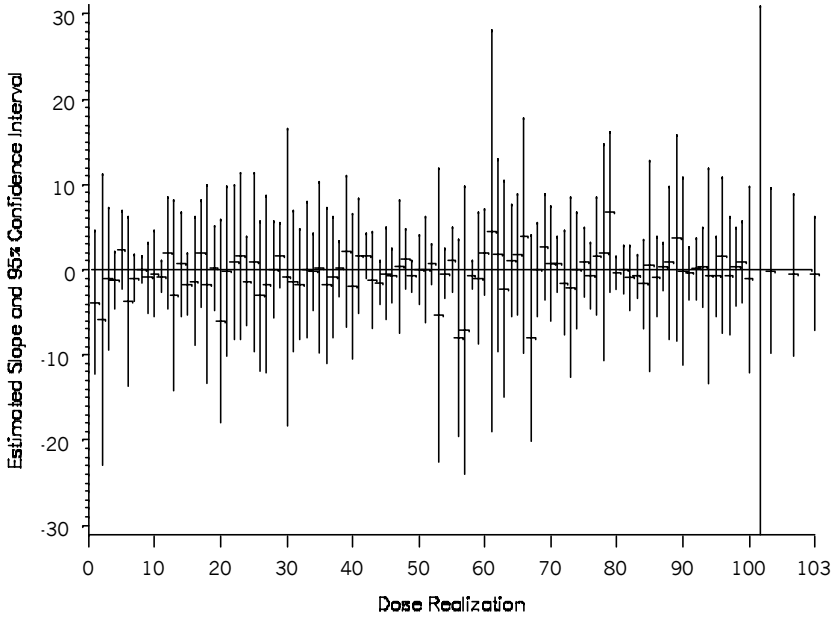
Parameter	Thyroid Mass (gm)
No. of living evaluable participants	3153
Estimated average background thyroid mass for women	7.69 (7.43, 7.96)
Estimated average background thyroid mass for men	11.51 (11.11, 11.92)
Estimated percentage change in average thyroid mass per Gy	–0.1% (–9.3%, +10.0%)
Statistical significance of dose-response (two-tailed p-value)	0.98

Entries for model parameters (background means and slope) in the tableare estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses.

Q.8.a. Uncertainty

The estimated slopes of the sex-stratified linear dose-response model for log of thyroid mass are shown in Figure IX.Q-22 for each of the 100 dose realizations produced by the CIDER model. The confidence intervals in that figure were calculated at the 98.33% confidence level, i.e., were adjusted by the Bonferroni technique for the simultaneous estimation of the slope and two sex-specific background rates. While the point estimate of the slope was greater than 0 for 45 of the 100 realizations, the confidence interval included 0 for all of the 100 realizations. Also shown in Figure IX.Q-22 (to the right of realization 100) are the estimates and confidence intervals calculated using (from left to right) the median, geometric mean, and mean of each participant's 100 dose realizations.

Figure IX.Q-22. Estimated Dose-Response for Thyroid Mass, by Dose Realization



R. Summary of Dose-Response Results

The primary evaluation of dose-response relationships focused on twelve categories of thyroid disease, hyperparathyroidism, and ultrasound-detected thyroid abnormalities of the thyroid. For each of these 14 outcome categories a primary case definition was specified based on the most definitive and valid diagnostic criteria available. The principal dose-response analysis used this primary definition of outcome, individual radiation dose estimates (the median for each individual) based on individual residence history, and on dietary consumption data from the CATI when available or on HEDR default values when CATI data were not available. The results from these analyses using the primary outcome definition constitute the principal findings of the HTDS. These results are summarized in Table IX.R-1 which shows that there are no significant dose-responses for the outcomes considered.

Table IX.R-1. Summary of Dose-Response Results for Thyroid Disease Outcomes

Thyroid Disease	Estimated Background Rates Female	Estimated Background Rates Male	Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
Thyroid Cancer	.006 ± .002 (.001, .011)	.002 ± .001 (0*, .005)	.002 ± .004 (< –.001, .017)	0.25
Benign thyroid nodule	.100 ± .008 (.081, .119)	.049 ± .006 (.034, .064)	–.008 ± .015 (< –.022, .041)	0.68
Total thyroid neoplasia	.011 ± .003 (.004, .018)	.006 ± .002 (.001, .012)	.001 ± .006 (< –.003, .022)	0.42
Any thyroid nodule	.112 ± .008 (.092, .132)	.053 ± .006 (.038, .068)	–.007 ± .016 (< –.023, .043)	0.65
Hypothyroidism	.118 ± .009 (.097, .139)	.037 ± .006 (.023, .050)	–.006 ± .019 (< –.016, .047)	0.61
Autoimmune thyroiditis	.239 ± .012 (.212, .267)	.133 ± .010 (.109, .156)	–.026 ± .026 (< –.057, .044)	0.82
Graves disease	.016 ± .004 (.008, .025)	.004 ± .002 (0*, .009)	–.001 ± .009 (< –.002, .024)	0.56
Autoimmune thyroid disease	.255 ± .012 (.227, .283)	.136 ± .010 (.112, .160)	–.024 ± .027 (< –.058, .048)	0.80
Hyperthyroidism	.077 ± .007 (.060, .094)	.015 ± .004 (.006, .025)	.011 ± .015 (< –.008, .052)	0.22
Multinodular thyroid gland	.040 ± .005 (.027, .053)	.014 ± .004 (.006, .023)	–.006 ± .016 (NE, .014)	0.88
Simple goiter	.006 ± .002 (.001, .011)	.003 ± .002 (0*, .008)	–.001 ± .008 (NE, .012)	0.74
Other thyroid disease	.010 ± .003 (.003, .016)	.003 ± .002 (0*, .008)	.002 ± .007 (< –.002, .024)	0.39

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, ">" indicates that the upper confidence limit is greater than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). "0*" indicates that the lower confidence limit for a background rate was less than 0.

Table IX.R-1. Summary of Dose-Response Results for Thyroid Disease Outcomes (continued)

Thyroid Disease	Estimated Background Rates Female	Estimated Background Rates Male	Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (one-tailed p-value)
Hyperparathyroidism	.006 ± .003 (0*, .013)	.001 ± .002 (0*, .006)	−.000 ± .018 (NE, .013)	0.61
Any UDA	.552 ± .014 (.519, .586)	.365 ± .014 (.332, .399)	.031 ± .038 (−.059, .116)	0.21
Palpable UDA	.090 ± .008 (.070, .110)	.043 ± .006 (.029, .057)	−.018 ± .023 (NE, .015)	0.95
Nonpalpable focal UDA	.451 ± .014 (.417, .484)	.303 ± .013 (.270, .335)	.027 ± .037 (−.061, .115)	0.23
Diffuse UDA	.174 ± .011 (.148, .199)	.084 ± .009 (.064, .105)	.029 ± .028 (−.029, .100)	0.14

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses ("<" indicates that the lower confidence limit is less than the indicated value, "NE" indicates the confidence limit was not estimated due to its close proximity to the point estimate). "0*" indicates that the lower confidence limit for a background rate was less than 0.

Less definitive criteria to identify cases were also defined for each outcome category using less definitive diagnostic criteria. Dose-response analyses were also conducted for each of these alternative definitions. In addition, dose-response analyses were conducted for six outcome categories based on the results of laboratory assays, and for thyroid mass estimated from the ultrasound scan (Table IX.R-2). The primary analysis for each outcome used the method of maximum likelihood to estimate the background rates or averages for women and men, and the slope of the sex-stratified linear models. Estimates of the parameters were also calculated using the method of least squares, once with doses treated as a continuous quantitative variable ("ungrouped analysis"), and again with doses treated as a categorical variable ("grouped analysis"). Linear quadratic and logistic dose-response models were also considered as alternatives to the linear model. Dose-response analyses for all outcomes were repeated using two alternative sets of individual dose estimates, and two alternative representations of exposure that did not use the HEDR models to estimate individual radiation dose. Efforts were also made to evaluate the influence of uncertainties in individual dose estimates on the fitted dose-response relationships for the primary case definition in each outcome category.

Table IX.R-2. Summary of Dose-Response Results for Laboratory Values and Thyroid Mass

Outcome	Estimated Background Rates Female	Estimated Background Rates Male	Estimated Slope of Dose-Response (per Gy)	Statistical Significance of Dose-Response (two-tailed p-value)
TSH by RIA (μIU/ml)	.90 ± .04 (.80, 1.00)	.89 ± .04 (.79, .99)	.02 ± .16 (–.36, .40)	0.90
TSH by EIA-1 (μIU/ml)	.46 ± .03 (.37, .54)	.31 ± .03 (.23, .39)	.13 ± .11 (–.13, .39)	0.22
TSH by EIA-2 (μIU/ml)	.39 ± .02 (.33, .45)	.23 ± .03 (.17, .29)	.01 ± .06 (–.13, .16)	0.82
T4 (μg/dl)	2.02 ± .01 (2.00, 2.03)	1.88 ± .01 (1.87, 1.90)	–.004 ± .02 (–.05, .04)	0.84
T3RU (μg/dl)	.021 ± .005 (.009, .032)	–.160 ± .005 (-.170, -.149)	–.01 ± .01 (–.04, .02)	0.36
FTI	1.998 ± .005 (1.99, 2.01)	2.044 ± .005 (2.03, 2.05)	.02 ± .01 (–.02, .05)	0.23
Anti-TPO (IU/ml)	–3.64 ± .42 (–4.65, 2.64)	–6.65 ± .54 (–7.95, -5.35)	–.39 ± .88 (–2.48, 1.71)	0.66
AMA (U/ml)	0.26 ± .30 (–.46, .99)	–1.24 ± .36 (–2.11, -.36)	–.51 ± .80 (–2.43, 1.42)	0.52
Anti-TG (IU/ml)	–4.01 ± .23 (–4.57, –3.45)	–5.71 ± .29 (–6.42, -5.01)	–.64 ± .50 (–1.84, 0.56)	0.20
Serum calcium (mg/dl)	9.2 ± .01 (9.14, 9.20)	9.2 ± .01 (9.16, 9.22)	–.09 ± .03 (–.16, –.01)	0.0074
Thyroid mass (gm)	2.04 ± .01 (2.00, 2.07)	2.44 ± .01 (2.41, 2.48)	–.00 ± .04 (–.10, .10)	0.98

Entries in the table are estimate ± standard error, with Bonferroni-adjusted 95% confidence interval in parentheses.

In overall summary of the dose-response results, there was no evidence of a statistically significant association between estimated thyroid radiation dose from Hanford and the cumulative incidence of any of the 14 primary thyroid or parathyroid disease outcomes or the prevalence of thyroid UDAs. There was also no evidence of any statistically significant dose-response relationship for any of the alternative definitions of outcome. The findings were essentially unchanged for analyses based on either of the two alternative sets of individual dose estimates. The results remained the same after taking into account (adjusting for the effects of) several factors that could potentially confound the relationship between radiation dose and the outcome of interest. There was no evidence of any statistically significant dose-response for any outcome that might be different from the linear model used in the primary analyses (e.g., a linear quadratic relationship). Incorporation of uncertainty in the dose estimates did not materially change the primary results for any of the outcomes.

The study also found no statistically significant associations between estimated thyroid dose from Hanford's ^{131}I and the average values of tests for thyroid function (TSH, T4, T3RU, FTI), of tests for anti-thyroid immune response (anti-TPO, AMA, anti-TG), or of thyroid mass. Only serum calcium, which was measured as a screening test for hyperparathyroidism, was found to vary significantly in relation to estimated thyroid dose from Hanford's ^{131}I: average calcium levels decreased significantly with increasing estimated thyroid radiation dose. However the decrease was small enough that calcium levels remained within the normal range, and less than 1% of the study participants were hypocalcemic.

Presented below are more detailed summaries of the results for each of the primary outcomes investigated.

Thyroid Cancer

Twenty (0.6%) of the 3440 living evaluable participants were diagnosed with thyroid cancer; 13 women (0.7%) and 7 men (0.4%). In all but one case, the diagnosis was based on histologic evidence from the HTDS examination (12) or prior histologic evidence (7).

Using the primary definition (19 total cases; 14 in-area) and maximum likelihood analysis of the sex-stratified linear probability model, the risk of thyroid cancer did not increase significantly with estimated dose ($p = 0.25$), with an estimated slope of 0.002 per Gy, and Bonferroni-adjusted 95% CI ranging from less than –0.001 to 0.017 per Gy. Results obtained by least squares analysis using ungrouped or grouped data were similar. There was no evidence from the linear-quadratic or logistic regression model that the cumulative incidence of thyroid cancer increased significantly with increasing dose. Analyses which considered less definitive criteria to identify cases and alternative dose estimates or representations of exposure revealed no evidence of a dose-response relationship. Incorporation of uncertainty in the dose estimates did not materially change the primary results.

Benign Thyroid Nodule

Two hundred and forty-nine (7.2%) of the 3440 living evaluable participants had a diagnosis of benign thyroid nodule based on histologic or cytologic evidence arising from the HTDS examination or from a prior diagnosis; 170 (9.7%) women and 79 (4.7%) men. An additional 38 (1.1%) participants had diagnoses classified as clinical, and another 10 (0.3%) had diagnoses based solely on a report by the participant or his/her CATI respondent.

Using the primary definition (249 total cases; 235 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the risk of benign thyroid nodule did not increase significantly with estimated dose ($p = 0.68$), with an estimated slope of -0.008 per Gy, and Bonferroni-adjusted 95% CI ranging from less than –0.022 to 0.041 per Gy. Results obtained by least squares analysis using ungrouped or grouped data were similar. There was no evidence from the linear-quadratic or logistic regression model that the cumulative incidence of benign thyroid nodule increased with increasing dose. Analyses which

considered less definitive criteria to identify cases, as well as other disease outcomes related to benign nodules (e.g., benign nodules and nodules suspicious for follicular neoplasm, benign nodule excluding non-neoplastic disease, solitary nodule detected without ultrasound, benign nodule excluding colloid-only nodules, and benign colloid nodules), and analyses which considered alternative dose estimates or representations of exposure, revealed no evidence of a dose-response relationship. Accounting for potential confounding or effect modification, and incorporation of uncertainty in the dose estimates, did not materially change the primary results.

Total Thyroid Neoplasia

This outcome was defined to include participants with thyroid cancer based on HTDS or prior histology or benign thyroid nodule with a histologic type of follicular adenoma, based on HTDS or prior histology. A total of 33 (1.0%) of the 3440 living evaluable participants were included in this category; 20 (1.1%) women and 13 (0.8%) men.

Using the primary definition (33 total cases; 28 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the risk of total thyroid neoplasia did not increase significantly with estimated dose ($p = 0.42$), with an estimated slope of 0.001 per Gy, and Bonferroni-adjusted 95% CI ranging from less than –0.003 to 0.022 per Gy. Results obtained by least squares analysis using ungrouped or grouped data were similar. There was no evidence from the linear-quadratic or logistic regression model that the cumulative incidence of total thyroid neoplasia increased with increasing dose. Analyses using alternative dose estimates or representations of exposure revealed no evidence of a dose-response relationship. Incorporation of uncertainty in the dose estimates, did not materially change the primary results.

Any Thyroid Nodule

This outcome was defined by the diagnosis of one or more of the following: benign thyroid nodule, thyroid cancer, or nodule suspicious for follicular neoplasm. A total of 281 (8.2%) of the 3440 living evaluable participants had this outcome based on histologic or cytologic evidence arising from the HTDS examination or from a prior diagnosis: 193 (11.0%) women and 88 (5.2%) men. Another 39 (1.1%) were based on clinical diagnoses by the HTDS or prior (palpable nodule with no available cytology or histology), and there were 10 living evaluable participants with a diagnosis of any thyroid nodule based solely on reports from the participant or his/her CATI respondent.

Using the primary definition (281 total cases; 261 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the risk of any thyroid nodule did not increase significantly with estimated dose ($p = 0.65$), with an estimated slope of –0.007 per Gy, and Bonferroni-adjusted 95% CI ranging from less than –0.023 to 0.043 per Gy. Results obtained by least squares analysis using ungrouped or grouped data were similar. There was no evidence from the linear-quadratic or logistic regression model that the cumulative incidence of any thyroid nodule increased with increasing dose. Analyses which considered less definitive criteria to identify cases and alternative dose estimates or representations of exposure revealed no evidence of a dose-response relationship. Accounting for potential confounding or effect modification, and incorporation of uncertainty in the dose estimates, did not materially change the primary results.

Hypothyroidism

Two hundred and sixty-seven (7.8%) of the 3440 living evaluable participants had a diagnosis of hypothyroidism based on the HTDS evaluation or on medical records with supporting documentation; 204 (11.7%) women and 63 (3.7%) men. An additional 105 (3.1%) living evaluable participants had a diagnosis of hypothyroidism based on medical records but without supporting documentation, and 30

(0.9%) were inferred from past or current thyroxine therapy. There were 193 (5.6%) cases based solely on reports of hypothyroidism from the participant or his/her CATI respondent.

Using the primary definition (267 total cases; 246 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the risk of hypothyroidism did not increase significantly with estimated dose ($p = 0.61$). With an estimated slope of −0.006 per Gy, and Bonferroni-adjusted 95% CI ranging from less than −0.016 to 0.047 per Gy. Similar results were obtained using the least squares analyses of grouped or ungrouped data. There was no evidence from the linear-quadratic or logistic regression model that the cumulative incidence of hypothyroidism increased with increasing dose. Analyses which considered less definitive criteria to identify cases, as well as permanent hypothyroidism, and analyses which considered alternative dose estimates or representations of exposure, revealed no evidence of a dose-response relationship, although the estimated regression coefficients from logistic regression analyses using less definitive criteria to identify cases were somewhat larger. Accounting for potential confounding or effect modification, and incorporation of uncertainty in the dose estimates, did not materially change the primary results.

Autoimmune (Hashimoto's) Thyroiditis

A total of 625 (18.2%) of the 3440 living evaluable participants had a diagnosis of autoimmune thyroiditis based on the HTDS evaluation or medical records with supporting documentation; 403 (23.1%) women and 222 (13.1%) men. Another three cases were based on medical records without supporting documentation, and one case was based solely on a report by the participant or his/her CATI respondent.

Using the primary definition (625 total cases; 582 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the risk of autoimmune thyroiditis did not increase significantly with estimated dose ($p = 0.82$), with an estimated slope of −0.026 per Gy, and Bonferroni-adjusted 95% CI ranging from less than −0.057 to 0.044 per Gy. Similar results were obtained when the least squares model was fit using ungrouped or grouped data. There was no evidence from the linear-quadratic or logistic regression model that the cumulative incidence of autoimmune thyroiditis increased with increasing dose. Analyses which considered less definitive criteria to identify cases, additional outcomes related to the assay for antithyroid immune response, and autoimmune thyroiditis in combination with non-iatrogenic, permanent hypothyroidism, as well as analyses which considered alternative dose estimates or representations of exposure, revealed no evidence of a dose-response relationship. Accounting for potential confounding or effect modification, and incorporation of uncertainty in the dose estimates, did not materially change the primary results.

Graves Disease

A total of 34 (1.0%) of the 3440 living evaluable participants had a diagnosis of Graves Disease based on the HTDS evaluation or on medical records with supporting documentation; 28 (1.6%) women and 6 (0.4%) men. Three (0.1%) living evaluable participants had a diagnosis of Graves Disease based on medical records without supporting documentation, and an additional thirteen (0.4%) were based solely on a report from the participant or his/her CATI respondent.

Using the primary definition (34 total cases; 32 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the risk of Graves Disease did not increase significantly with estimated dose ($p = 0.56$), with an estimated slope of −0.001 per Gy, and Bonferroni-adjusted 95% CI ranging from less than −0.002 to 0.024 per Gy. Results obtained by least squares analysis using ungrouped or grouped data were similar. There was no evidence from the linear-quadratic or logistic regression model that the cumulative incidence of Graves Disease increased with increasing dose. Analyses which considered less definitive criteria to identify cases and alternative dose estimates or representations of exposure revealed no evidence of a dose-response relationship. Accounting for potential confounding or

effect modification, and incorporation of uncertainty in the dose estimates, did not materially change the primary results.

Autoimmune Thyroid Disease

Autoimmune thyroid disease was defined by a diagnosis of autoimmune (Hashimoto's) thyroiditis or Graves disease based on the HTDS evaluation or medical records with supporting documentation. A total of 659 (19.2%) of the 3440 living evaluable participants were included in this category; 431 (24.7%) women and 228 (13.5%) men. These included 625 with autoimmune (Hashimoto's) thyroiditis and 34 others with diagnoses of Graves disease. An additional 4 (0.1%) living evaluable participants had a diagnosis of autoimmune thyroid disease based on medical records without supporting documentation (three with autoimmune thyroiditis, one with Graves disease). Eleven others (0.3%) were based solely on a report by the participant or his/her CATI respondent (one with autoimmune thyroiditis, 10 with Graves disease).

Using the primary definition (659 total cases; 614 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the risk of autoimmune thyroid disease did not increase significantly with estimated dose ($p = 0.80$), with an estimated slope of –0.024 per Gy, and Bonferroni-adjusted 95% CI ranging from less than –0.058 to 0.048 per Gy. Similar results were obtained when the least squares model was fit using ungrouped or grouped data. There was no evidence from the linear-quadratic or logistic regression model that the cumulative incidence of autoimmune thyroid disease increased with increasing dose. Analyses which considered less definitive criteria to identify cases and alternative dose estimates or representations of exposure revealed no evidence of a dose-response relationship. Accounting for potential confounding or effect modification, and incorporation of uncertainty in the dose estimates, did not materially change the primary results.

Hyperthyroidism

A total of 161 (4.7%) of the 3440 living evaluable participants were diagnosed with hyperthyroidism based on the HTDS evaluation or medical records with supporting documentation; 134 (7.7%) women and 27 (1.6%) men. An additional 14 (0.4%) living evaluable participants had a diagnosis of hyperthyroidism based on medical records without supporting documentation, and 21 (0.6%) were based solely on a report from the participant or his/her CATI respondent. It is important to note that these 196 cases included a substantial number of iatrogenic cases (caused by excess thyroid hormone replacement). Since endogenous hyperthyroidism was of particular importance, analyses that focused on cases of non-iatrogenic hyperthyroidism were emphasized in this study.

Using the primary definition (161 total cases; 155 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the risk of hyperthyroidism did not increase significantly with estimated dose ($p = 0.22$), with an estimated slope of 0.011 per Gy, and Bonferroni-adjusted 95% CI ranging from less than –0.008 to 0.052 per Gy. Similar results were obtained when the least squares model was fit using ungrouped or grouped data. There was no evidence from the linear-quadratic or logistic regression model that the cumulative incidence of hyperthyroidism increased with increasing dose. Analyses which considered less definitive criteria to identify cases, as well as non-iatrogenic hyperthyroidism, and analyses which considered alternative dose estimates or representations of exposure, revealed no evidence of a dose-response relationship. Accounting for potential confounding or effect modification, and incorporation of uncertainty in the dose estimates, did not materially change the primary results.

Multinodular Thyroid Gland

A total of 95 (2.8%) of the 3440 living evaluable participants had a diagnosis of multinodular thyroid gland based on the HTDS evaluation; 73 (4.2 %) women and 22 (1.3 %) men. An additional

nineteen (0.6%) living evaluable participants had a diagnosis of multinodular thyroid gland based on medical records, and one diagnosis was based solely on a report from the participant or his/her CATI respondent.

Using the primary definition (95 total cases; 85 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the risk of multinodular thyroid gland did not increase significantly with estimated dose ($p = 0.88$), with an estimated slope of -0.006 per Gy. The lower limit of the Bonferroni-adjusted 95% confidence interval was not estimated, but the upper limit was 0.014 per Gy. When the model was fit by the method of least squares, the estimated slope using either ungrouped or grouped data was even more negative than the maximum likelihood estimate, thereby providing no evidence that risk of multinodular gland increased with increasing dose ($p = 0.89$ and 0.83, respectively). There was no evidence from the linear-quadratic or logistic regression model that the cumulative incidence of multinodular thyroid gland increased with increasing dose. Analyses which considered less definitive criteria to identify cases and alternative dose estimates or representations of exposure revealed no evidence of a dose-response relationship. Accounting for potential confounding or effect modification, and incorporation of uncertainty in the dose estimates, did not materially change the primary results.

Simple Goiter

The diagnosis of simple goiter was uncommon, with only 14 (0.4%) of the 3440 living evaluable participants having this diagnosis based on HTDS evaluation; 9 (0.5%) women and 5 (0.3%) men. Another 28 (0.8%) had diagnoses based on medical records, and for an additional 28 (0.8%) the diagnosis was based solely on a report by the participant or his/her CATI respondent.

Using the primary definition (14 total cases; all in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the risk of simple goiter did not increase significantly with estimated dose ($p = 0.74$), with an estimated slope of -0.001 per Gy. The lower limit of the Bonferroni-adjusted 95% confidence interval was not estimated, but the upper limit was 0.012 per Gy. When the model was fit by the method of least squares, the estimated slope using either ungrouped or grouped data was even more negative than the maximum likelihood estimate, thereby providing no evidence that risk of simple goiter increased with increasing dose ($p = 0.79$ and 0.70, respectively). There was no evidence from the linear-quadratic or logistic regression model that the cumulative incidence of simple goiter increased with increasing dose. Analyses which considered less definitive criteria to identify cases and alternative dose estimates or representations of exposure revealed no evidence of a dose-response relationship. Incorporation of uncertainty in the dose estimates, did not materially change the primary results.

Other Thyroid Disease

Four living evaluable participants, all in the in-area group, had diagnoses of other thyroid disease based on their HTDS examinations or medical records with supporting documentation. These included two cases of subacute thyroiditis in women; one case of familial thyroglobulin binding deficiency in a male; and one case of secondary hypothyroidism in a female. The first alternative definition added only two cases with diagnoses based on medical records without supporting documentation. Both were cases of subacute thyroiditis in women. For both the primary and first alternative definition of other thyroid disease, there were too few cases for meaningful estimation of the radiation dose-response.

The second alternative definition added 20 participants, primarily with participant or CATI respondent reports of past thyroid disease of unknown type. This brought the total number of cases to 26, of whom four were out-of-area participants. Based on maximum likelihood analysis of the sex-stratified linear probability model using this case definition, the estimated slope was slightly greater than zero (0.002 per Gy) with Bonferroni-adjusted 95% CI ranging from less than –0.002 to 0.024 per Gy, providing no evidence that cumulative incidence increased significantly with increasing dose (one-tailed $p = 0.39$).

Because the number of cases in this category was small, and the diagnoses were heterogeneous and mostly unknown, further analyses of this outcome were not performed.

Hyperparathyroidism

A total of 12 (0.3%) living evaluable participants had a diagnosis of hyperparathyroidism based on the HTDS evaluation or on medical records with supporting documentation; 10 (0.6 %) women and 2 (0.1%) men. Another two diagnoses were based on a report from the participant or his/her CATI respondent. One additional living evaluable participant who did not meet the study's criteria for hyperparathyroidism nevertheless had an elevated calcium in the presence of a high normal PTH level, when the PTH should have been suppressed, highly suggestive of hyperparathyroidism. This participant was included as a case in an additional analysis.

Using the primary definition (12 total cases; 11 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the risk of hyperparathyroidism did not increase significantly with estimated dose ($p = 0.61$), with an estimated slope of –0.0001 per Gy. The lower limit of the Bonferroni-adjusted 95% confidence interval was not estimated, but the upper limit was 0.013 per Gy. When the model was fit by the method of least squares, the estimated slope using either ungrouped or grouped data was slightly more negative than the maximum likelihood estimate, thereby providing no evidence that risk of hyperparathyroidism increased with increasing dose ($p = 0.74$ and 0.75, respectively). There was no evidence from the linear-quadratic or logistic regression model that the cumulative incidence of any thyroid nodule increased with increasing dose. Analyses which considered less definitive criteria to identify cases and alternative dose estimates or representations of exposure revealed no evidence of a dose-response relationship. Incorporation of uncertainty in the dose estimates, did not materially change the primary results.

Ultrasound-Detected Abnormalities of the Thyroid (Thyroid UDAs)

The thyroid gland was visible in the ultrasound examinations of 3429 of the 3440 living evaluable participants. For 11 participants the thyroid was not visible, 10 because of thyroid surgery and one because the sonographer couldn't adequately visualize the thyroid. Among the 3429 whose thyroids were visible, 1596 (46.5%) had one or more ultrasound-detected thyroid abnormalities (thyroid UDAs); 964 (55.5 %) women and 632 (37.4 %) men. Ultrasound findings were categorized as palpable thyroid UDAs (224 or 6.5%), nonpalpable focal thyroid UDAs (1309 or 38.2%), and diffuse thyroid UDAs (458 or 13.4%). All three types of UDA were more frequent among women than men. Ultrasound-detected thyroid abnormalities were based only on the HTDS evaluation, not on any prior ultrasounds.

For any UDA (1596 total cases; 1481 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the risk of any UDA did not increase significantly with estimated dose ($p = 0.21$), with an estimated slope of 0.031 per Gy, and Bonferroni-adjusted 95% CI ranging from –0.059 to 0.116 per Gy. Estimation by least squares using the ungrouped data gave nearly identical results, and the least squares fit to the grouped data were similar. There was no evidence from the linear-quadratic or logistic regression model that the cumulative incidence of any UDA increased with increasing dose. Analyses which considered alternative dose estimates or representations of exposure revealed no evidence of a dose-response relationship. Accounting for potential confounding or effect modification, and incorporation of uncertainty in the dose estimates, did not materially change the primary results.

For palpable UDAs (224 total cases; 204 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the risk of a palpable UDA did not increase significantly with estimated dose ($p = 0.95$), with an estimated slope of -0.018 per Gy. The Bonferroni-adjusted lower 95% confidence limit was not estimated due to the magnitude of the negative slope estimate, however the upper confidence limit was 0.015 per Gy. Estimation by least squares using either the ungrouped or grouped data gave nearly identical results. There was no evidence from the linear-quadratic or logistic regression model that the prevalence of palpable thyroid UDAs increased with increasing dose. Analyses which considered

alternative dose estimates or representations of exposure revealed no evidence of a dose-response relationship. Accounting for potential confounding or effect modification, and incorporation of uncertainty in the dose estimates, did not materially change the primary results.

For nonpalpable focal UDA (1309 total cases; 1217 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the risk of a nonpalpable focal UDA did not increase significantly with estimated dose ($p = 0.23$), with an estimated slope of 0.027 per Gy, and Bonferroni-adjusted 95% CI ranging from –0.061 to 0.115 per Gy. Estimation by least squares using either the ungrouped or grouped data gave nearly identical results. There was no evidence from the linear-quadratic or logistic regression model that the prevalence of nonpalpable focal thyroid UDAs increased with increasing dose. Analyses which considered alternative dose estimates or representations of exposure revealed no evidence of a dose-response relationship. Accounting for potential confounding or effect modification, and incorporation of uncertainty in the dose estimates, did not materially change the primary results.

To assess whether the dose-response results might be affected by the size of focal thyroid UDAs, three additional outcomes were analyzed. These included the presence of a focal UDA with maximum dimension at least 5 mm, the presence of a focal UDA with maximum dimension at least 10 mm, and the presence of a focal UDA with average dimension at least 15 mm. These additional analyses applied only to palpable and nonpalpable focal thyroid UDAs, since diffuse UDAs were not defined by any size criterion. In none of these additional analyses was there any evidence that the risk of having a focal UDA of a particular size increased with increasing dose ($p=0.64$, 0.88 and 0.53 for the presence of focal UDA with maximum dimension of 5 mm, maximum dimension of 10 mm and average dimension of 15 mm, respectively).

Additional analyses were performed to investigate whether the number of thyroid UDAs detected in individual participants might increase in relation to estimated thyroid radiation dose. For each living evaluable participant with an HTDS ultrasound examination, the numbers of focal thyroid UDAs with maximum dimension ≥ 5 mm, maximum dimension ≥ 10 mm, and average dimension ≥ 15 mm were counted. Study participants had as many as nine thyroid UDAs with maximum dimension ≥ 5 mm, although the majority (60% of the women and 74% of the men) had no such thyroid UDAs. The overall average number of thyroid UDAs of this size was 0.84 per person for women, and 0.47 per person for men. Results of fitting sex-stratified Poisson regression models for the relationship between estimated thyroid radiation dose and number of focal thyroid UDAs indicated that the average number of such thyroid UDAs per person did not increase significantly with estimated dose ($p = 0.80$, 0.48 and 0.43 for the number of thyroid UDAs with maximum dimension of 5mm, maximum dimension of 10 mm and average dimension of 15 mm, respectively.). The Bonferroni-adjusted 95% confidence interval for the dose-response parameter for number of thyroid UDAs with maximum dimension of 5 mm ranged from 0.72 to 1.17, encompassing a range from 28% decrease to 17% increase per Gy. The results for the number of thyroid UDAs with maximum dimension of 10 mm and average dimension of 15 mm were similar.

Using the primary definition of diffuse UDA (458 total cases; 428 in-area), and maximum likelihood analysis of the sex-stratified linear probability model, the risk of a diffuse UDA did not increase significantly with estimated dose ($p = 0.14$), with an estimated slope of 0.029 per Gy, and Bonferroni-adjusted 95% CI ranging from –0.029 to 0.100 per Gy. Estimation by least squares using either the ungrouped or grouped data gave nearly identical results. There was no evidence from the linear-quadratic or logistic regression model that the prevalence of diffuse thyroid UDAs increased with increasing dose. Analyses which considered alternative dose estimates or representations of exposure revealed no evidence of a dose-response relationship. Accounting for potential confounding or effect modification, and incorporation of uncertainty in the dose estimates, did not materially change the primary results.

Laboratory Tests

Of the 3191 living evaluable in-area participants, 3183 (99.7%) consented to provide a blood specimen at their HTDS clinic. Several laboratory assays were conducted to evaluate thyroid function, anti-thyroid antibody response, and serum calcium level. In addition to the dose-response analyses conducted of specific thyroid disease outcomes which incorporated information from these tests in the determination of the diagnosis, dose-response analyses were also conducted to investigate whether there were associations between the laboratory values from these tests and estimated thyroid radiation dose from Hanford (i.e., regardless of thyroid disease diagnosis).

Thyroid stimulating hormone (TSH) levels were determined according to three different tests over the course of the study. Of the 3183 living evaluable in-area participants who provided blood samples, 222 were receiving exogenous thyroid hormone at the time of their HTDS clinic. These 222 were excluded from the analyses of TSH. Among the remaining 2961 living evaluable participants, 584 had TSH measured by RIA, 810 by EIA-1, and 1567 by EIA-2. There was no evidence of a significant trend in relation to estimated radiation dose for any of the three assays considered individually ($p = 0.90$, 0.22, and 0.82 for RIA, EIA-1, and EIA-2, respectively). When a generalization of the sex-stratified model was examined using all 2961 living evaluable participants with TSH measurements, in which the mean values of log(TSH) were assumed to differ between the sexes and to differ according to the type of assay, there was still no significant trend of average log(TSH) in relation to estimated thyroid radiation dose ($p = 0.42$).

Analyses of total thyroxine (T4) and triiodothyronine resin uptake (T3RU) were also conducted excluding the 222 living evaluable in-area participants who were receiving exogenous thyroid hormone at the time of their HTDS clinic. Additionally, the T4 and T3RU values were unknown for two in-area participants due to insufficient volumes of collected blood. There was no significant trend of either T4 or T3RU in relation to estimated radiation dose ($p = 0.84$ and 0.36, respectively).

Free thyroxine index (FTI) was analyzed excluding the 222 living evaluable in-area participants who were receiving exogenous thyroid hormone at the time of their HTDS clinic. The FTI values were unknown for six additional in-area participants: the two with unknown T4 and T3RU, and four others for whom either T4 or T3RU was below its level of detection. There was no significant trend of FTI in relation to estimated radiation dose ($p = 0.23$).

Anti-TPO or AMA values were used for the anti-thyroid autoimmune response evaluations of 1562 and 1620 in-area living evaluable participants, respectively. Neither assay result was available for eight participants who declined to provide a blood sample, and for one other whose sample was of insufficient volume. There was no significant trend of either assay result in relation to estimated radiation dose ($p = 0.66$ for anti-TPO, 0.52 for AMA). Anti-thyroglobulin antibody (anti-TG) values were available for 3170 of the in-area living evaluable participants. There was no significant trend of anti-TG in relation to estimated radiation dose (two-tailed $p = 0.20$).

Serum calcium levels were measured in an effort to identify study participants with hypercalcemia which might be secondary to hyperparathyroidism. Of the 3183 living evaluable in-area participants who provided blood samples, 227 with diagnoses of hypothyroidism or hyperthyroidism based on the HTDS examination were excluded from the primary analysis of serum calcium levels. Two additional participants did not have serum calcium data due to insufficient volumes of collected blood. There was a statistically significant trend of decreasing serum calcium level in relation to increasing radiation dose ($p = 0.0074$), with an estimated slope of -0.09 per Gy and Bonferroni-adjusted 95% CI ranging from -0.16 to -0.01. Although there is no readily apparent explanation for this result, this finding deserves further comment. First, it should be noted that the laboratory test used measured the total serum calcium and not ionized calcium, which is the true measure of normal calcium levels in the blood. Thus, it cannot be certain that a dose effect would be present if ionized calcium rather than total calcium had been measured. Second, the outcome for which calcium was being measured, hyperparathyroidism, was not found to be associated with radiation dose. Third, the dose effect occurred primarily in the normal calcium range. For both women and men, the estimated background means were about $9.2 \pm .01$, consistent with the normal range of the test

(8.4-10.2). Only 0.9% of the cohort had low calcium levels less than 8.4 mg/dl (hypocalcemia). There was no statistically significant relationship between hypocalcemia and radiation dose. Even at a dose of 3 Gy (3000 mGy), which is larger than the maximum estimated dose of any study participant, average calcium levels predicted by the regression model were will within the normal range. Therefore, despite the *statistically* significant decrease in calcium levels with increasing dose, the resulting effect or clinical impact does not appear to be *clinically* significant.

Estimates of thyroid mass were available for 3400 living evaluable participants for whom both lobes of the thyroid were visible on ultrasound; 3153 were in-area participants. There was no significant trend of thyroid mass in relation to estimated radiation dose ($p = 0.98$).

X. DISCUSSION

A. Summary of Study Design and Execution

The purpose of the Hanford Thyroid Disease Study was to determine whether thyroid morbidity is increased among people exposed to atmospheric releases of ^{131}I from the Hanford Nuclear Site between 1944 and 1957. The primary objective of the research was to describe in what way any increase in thyroid disease observed is related to the dose of radiation received (that is, to describe the characteristics of any dose-response relationship). Additional objectives of the project were to: 1) determine whether hyperparathyroidism is increased among people exposed to the Hanford radiation releases; 2) assess the methods used to carry out such a study, and the degree to which such an investigation can be successfully planned and executed; and 3) provide information to residents of the surrounding communities regarding the conduct of the study and the findings and results.

In order to achieve the primary objective stated above, the study was conducted as a retrospective follow-up (cohort) study. As described more fully in section IV.A above, this design entailed the selection of a relatively large cohort of people who would have been exposed to Hanford radiation as young children, and who would represent the full range of possible doses to the thyroid from Hanford. The overall goal of the study was to locate all individuals in the cohort, obtain their consent to participate in the project, collect detailed information regarding their early childhood in order to estimate the dose of radiation they received to the thyroid from Hanford, and determine whether they have developed any form of thyroid disease or hyperparathyroidism since their exposure. The primary determination of whether there has been an increase in thyroid disease as a result of radiation exposure from Hanford was made by assessing the cumulative incidence of thyroid disease in relation to the level of individual thyroid radiation dose (i.e., the radiation dose-response), within members of the cohort.

This approach of using one population comprised of individuals with different levels of exposure to radiation has been used extensively in assessing the effects of radiation exposure in human populations. It is a common design in epidemiology, and has been of particular value in studies of atomic bomb survivors in Japan, in numerous studies of people exposed to radiation through medical procedures, and in the study of people exposed to radiation from atmospheric testing in Utah. This method is superior to the alternative approach of attempting to compare thyroid disease occurrence in a cohort under extensive study such as the HTDS cohort with that in a separate population presumed to be unexposed to radiation. This is because thyroid disease rates may be a function of a number of factors other than exposure to radiation, which may differ considerably between different populations. This is particularly true if one population is under careful study and diagnostic evaluation. Such differences can include: 1) the methods of diagnosis employed; 2) the extent to which diagnostic tests are implemented in a population (i.e., the thoroughness of the diagnostic process); 3) the dietary practices of the population; 4) the level of iodine in the diet; and 5) the composition of the population according to age, gender, and ethnicity. To the extent differences in such factors exist, it would be impossible to attribute any differences in thyroid disease rates observed to Hanford radiation exposure, as opposed to one or more of these other factors. The approach used in the HTDS is also superior to one which would implement the full HTDS protocol in a population geographically removed from the Hanford site, in an attempt to include people with no exposure from Hanford radiation. Although the methods and thoroughness of the diagnostic evaluation would be comparable under such circumstances, it would still not be possible to ensure comparability between the two study populations regarding the other types of possible differences listed above that could influence thyroid disease occurrence. Thus, to ensure as much comparability as possible regarding factors other than radiation that can influence the occurrence of thyroid disease, all comparisons of thyroid disease rates in relation to thyroid radiation dose level were made within the defined cohort.

The study cohort was defined based on place and year of birth in a manner designed to identify people with a full range of possible thyroid doses from Hanford. This was difficult to do, as no information was available regarding exposure or estimated dose to specific individuals. Using preliminary information available at the time of the design of the HTDS from the HEDR Project regarding the timing of the

radiation releases, movement of radioactive materials through the environment, and uptake by humans, a sampling scheme was developed to select individuals born in one of seven counties in the region around Hanford between 1940 and 1946.

From the outset it was recognized that such a design would present substantial challenges. The study would require that: 1) people be traced and located over a period of more than 40 years, based only on information contained on their birth certificate; 2) a person knowledgeable of each participant's early childhood be located, and be willing and able to provide detailed information about the person's childhood residence history and dietary habits; 3) participants be willing to travel to the Northwest and undergo a thorough medical evaluation for thyroid disease, including an ultrasound scan, blood tests, and potentially a thyroid biopsy; and 4) participants be willing to provide consent for independent review of prior medical records and diagnostic evaluations. Since no study of this type had ever been attempted before, it was not clear that such an approach would prove feasible.

As described more fully in section V above, the field components of the study were highly successful. A roster of 12,706 births was constructed from Washington State birth certificates, and 5199 individuals were selected for inclusion in the cohort. Of these, 4350 individuals (84%) were located alive and their identity confirmed, and 527 (10%) were confirmed as deceased. Importantly, success in locating individuals did not vary appreciably by sex, geographic region at the time of birth, or year of birth. Of those known to be deceased or who died prior to participation (16), a death certificate was obtained for verification of cause of death for 504 (93%).

Once contacted, individuals were cooperative and interested in participating. Of 4239 people contacted by phone to request participation, 3564 (84%) agreed to participate. Only 634 (15%) refused. Forty-one living located cohort members (0.9%) were determined to be unable to fully participate and were consequently not included in the study regardless of willingness to participate. Agreement to participate did not vary appreciably according to sex, geographic region at the time of birth, year of birth, or location of current residence. Of the 3564 who agreed to participate, 2712 participants identified a possible CATI respondent. Interviews were completed for 2266 (84%). However, not all of those for whom a CATI was completed attended a clinic. Thus, of the 3440 living evaluable participants included in the analysis, 2123 (62%) had a CATI interview that was used as the basis for dose estimation. Quality assessment of the respondent's ability to answer the interview questions was also performed (see section V.D). Following each section of the interview, interviewers recorded their assessment of how reliable the responses were for questions within that section. Assessments were recorded as: 1) Very Reliable; 2) Somewhat Reliable, or 3) Unreliable. Overall, the interviewers rated the quality of the data obtained in the CATI as very reliable.

It proved feasible to hold all clinics in the Pacific Northwest (all but one site was located in Washington), and participants were willing to travel from throughout the United States and even from abroad to attend clinics. Of those who agreed to participate, and who did not withdraw from the study at a later time, 97% (3447 of 3564) attended a HTDS clinic. Success in scheduling people for clinics did not vary substantially by sex, geographic area of birth, year of birth, or even current residence. Of those attending a clinic, almost all (> 99 %) participated in all aspects of the evaluation: In-Person Interview, thyroid ultrasound, blood tests, and clinical examination. Of the 272 participants for whom a fine needle aspiration biopsy of the thyroid was recommended, 259 (95%) underwent the procedure.

It also proved feasible to locate and retrieve prior medical records and materials. Attempting to locate and obtain records from as long ago as fifty years was expected to be one of the most difficult aspects of the HTDS. A total of 694 participants identified prior medical records of potential interest, and provided consent to the HTDS to request 1259 separate medical records. Of these, 795 (63%) were received by the HTDS from 494 of the 694 participants (71%). Pathology or cytology slides were requested for 52 of the 694 individuals identifying historical material. Of these, slides were received from 42 (81%).

The results of the field components of the HTDS reflect a relatively uniform and high level of success in achieving the objectives set forth for each. It proved feasible to identify a large group of people exposed at varying levels to radiation releases from Hanford, to locate and contact them, to enroll them in the study, to collect information needed to estimate their individual radiation dose to the thyroid, to

examine them for the presence or history of thyroid disease, and to review prior medical records relevant to prior evaluations or diagnoses of thyroid disease. Given the eligibility criteria for inclusion in this study, there was no evidence that success in these various tasks varied appreciably according to a person's sex, where they were born, when they were born, or where they currently live. Because these operational results provide no indication of substantial differential success in aspects of the study related to subject selection, inclusion, and data collection that might potentially bias or influence dose-response relationships, they provide an important framework for interpreting the specific findings regarding radiation dose and the thyroid outcomes under study.

B. Summary of Dose-Response Results

The primary evaluation of dose-response relationships focused on twelve categories of thyroid disease, ultrasound-detected abnormalities of the thyroid, and hyperparathyroidism. For each of these 14 outcome categories a primary case definition was specified based on the most definitive and valid diagnostic criteria available. The diagnostic information used for each primary outcome definition was obtained at the time of the participant's clinical evaluation at an HTDS clinic site. This information included results from thyroid physical examinations, laboratory tests, ultrasound scans, and thyroid biopsy results. For most outcomes, if a participant's prior medical records confirmed a diagnosis with the same diagnostic methods as those used at the HTDS clinic evaluation, such information was classified as having met the criteria for the primary outcome definition. The principal dose-response analysis used this primary definition of outcome, individual radiation dose estimates (the median for each individual) based on individual residence history, dietary consumption data from the CATI or expanded In-Person Interview when available, and HEDR default values when such data were not available. The results from these analyses using the primary outcome definition constitute the principal findings of the HTDS. The primary analysis for each outcome used the method of maximum likelihood to estimate the background rates or averages for women and men, and the slope of the sex-stratified linear models. Estimates of the parameters were also calculated using the method of least squares, once with doses treated as a continuous quantitative variable ("ungrouped analysis"), and again with doses treated as a categorical variable ("grouped analysis"). Linear quadratic and logistic dose-response models were also considered as alternatives to the linear model.

Alternative case definitions were also specified for each outcome category using less definitive diagnostic criteria. The diagnostic information used for the alternative case definitions did not meet the HTDS primary outcome criteria, but was obtained from additional sources. These sources included statements from medical records for which the diagnosis could not be confirmed, or reports from the participant or his or her CATI respondent of a diagnosis for which no medical records could be found. Dose-response analyses were also conducted for each of these alternative definitions. In those instances where an alternative definition resulted in a substantially greater number of people in the analysis than the primary definition, the dose-response results for the alternative definition are also presented in the Results section. In addition, dose-response analyses were conducted for six outcome categories based on the results of laboratory assays, and for thyroid mass estimated from the ultrasound scan.

All dose-response analyses for all outcome definitions were repeated using two alternative sets of individual dose estimates: 1) individual residence history, and only HEDR default data regarding dietary consumption (i.e., no data from the CATI or expanded In-Person Interview); and 2)individual residence history, dietary consumption data from the CATI or expanded In-Person Interview when available, and default values based on the HTDS CATI data when such individual data were not available (with the exception of consumption other than milk for expanded IPIs for which HEDR defaults were used). Further, two alternative representations of exposure were defined which did not use the HEDR models to estimate individual radiation dose. Although these categorizations of exposure were more crude than the individual quantitative estimates of dose, such analyses were performed as an alternative means of investigating a possible relationship between the thyroid outcomes under study and exposure to Hanford radiation that would be independent of the HEDR models and assumptions.

Because the primary focus of the HTDS analysis was to investigate possible dose-response relationships, and because individual radiation doses estimated in this study were characterized by some degree of uncertainty due to the uncertain nature of many of the parameters that determine dose, efforts were also made to evaluate the influence of dose uncertainties on the fitted dose-response relationships for the primary case definition in each outcome category. Two different approaches were used. First, the linear dose-response models were fit using each of the 100 realizations of thyroid dose obtained from the HEDR models. The estimate of, and 95% confidence interval for, the slope of the dose-response from each of these 100 analyses were displayed graphically to illustrate how the estimated radiation effect varied among the 100 realizations of dose, and how these estimates compared to the results based on the median or other average dose estimate.

In the second approach, Bayesian analysis was used to estimate the parameters of the logistic dose-response model that were adjusted for the effect of the dose uncertainty. This approach used the Gibbs sampling technique to estimate the marginal posterior distribution of the model parameters, conditional on the observed data, from the joint conditional distribution of the parameters and unobserved true doses. For each primary outcome definition, the estimated marginal posterior distribution of the regression coefficient was displayed. Also the median and appropriate percentiles of that distribution were used to derive uncertainty-adjusted point and confidence interval estimates of the dose-response coefficient, for comparison to the corresponding unadjusted estimates. As expected, the effect of this adjustment for uncertainty was generally to increase the magnitude of the estimated dose-response coefficient. That is, if the unadjusted estimate was less than 0, then the adjusted estimate was even more negative. Similarly, if the unadjusted estimate was greater than zero, the adjusted estimate was even larger. Also as expected, however, the adjustment for uncertainty reduced the precision with which the regression coefficient was estimated, i.e., the uncertainty-adjusted confidence intervals were wider than the corresponding unadjusted intervals. Consequently for none of the outcomes did the adjustment for dose uncertainty reveal a significant dose-response that was obscured in the unadjusted analyses.

In overall summary of the dose-response results, there was no evidence of a statistically significant association between estimated thyroid radiation dose from Hanford and the cumulative incidence of any of the 14 primary outcomes. There was also no evidence of any statistically significant dose-response relationship for any of the alternative definitions of outcome. These results were remarkably uniform. The findings were essentially unchanged for analyses based on either of the two alternative sets of individual dose estimates. The results remained the same after taking into account (adjusting for the effects of) several factors that could potentially confound the relationship between radiation dose and the outcome of interest. There was no evidence of any statistically significant dose-response for any outcome that might be different from the linear model used in the primary analyses (e.g., a linear-quadratic relationship). Incorporation of uncertainty in the dose estimates did not materially change the primary results for any of the outcomes.

C. Consideration of Factors Related to Study Design and Execution

In interpreting the findings of an epidemiologic study like the HTDS, it is important to consider the possible influence on the results of factors other than those directly accounted for in the analysis. Of particular concern is the possibility that the results could be due in part (or entirely) to artifacts or flaws in either the design or conduct of the study. A number of different factors are considered below in an attempt to better understand the absence of any dose-response relationships found with any of the outcomes investigated in this study.

C.1. Factors Related to Cohort Definition and Selection

A fundamental consideration in interpreting these results is the adequacy and appropriateness of the study group upon which all analyses are based. Two principal aspects of this question must be addressed: 1) was the definition and selection of the study group adequate in order to achieve the primary research objective; and 2) were the analyses based on an unbiased representation of this group?

The primary research objective of this study was to determine whether thyroid disease is increased among people exposed to atmospheric releases of radioactive iodine from the Hanford Nuclear Site between 1944 and 1957. A study group was defined with the intention of identifying individuals who could have been exposed based on their proximity to the site during the times when the releases were highest. Further, to have the greatest likelihood of detecting an effect of exposure, the study group was restricted to include people who would have been young children at the time of greatest exposure. This was based on the assumption that young children receive a higher dose to the thyroid from ^{131}I for the same level of exposure than do adolescents or adults, and that the thyroid gland in young children may be more susceptible to the effects of a given dose than in older people.

A study group was therefore defined based on births that occurred in the region. A roster of births to mothers living in a seven-county area between 1940 and 1946 was constructed from Washington State birth records. The HTDS cohort was selected from this roster using a stratified random sampling technique. Thus, the study group of 5199 individuals selected for inclusion in the HTDS reflects a random sample of a complete listing of births, and as such provides a population of people who could have been exposed to the Hanford releases at young ages. It does not define the total population in the region who could have been exposed, nor even the total population of the age range encompassed by the selected birth cohort who could have been exposed. However, in order to achieve the objectives set forth using a cohort study design, it is not necessary to investigate the entire population at risk or even a representative sample thereof. In this instance, restricting the definition of the cohort in some manner (in this case as a birth cohort) is not of concern in terms of introducing a possible bias in the dose-response. The more important issue is whether the study group is defined in a manner that will include people representing the entire range of possible doses, that will include adequate numbers of people with the highest as well as the lowest doses, and that will allow for uniform and complete follow-up to ascertain thyroid disease status for everyone in the group in the same manner. Based on what is known about the Hanford radiation releases and possible exposures to people in the region, the definition and selection of the HTDS cohort should be quite adequate to achieve the primary research objectives.

A more critical consideration is the question of who actually ended up participating in the study and contributing to the analyses of dose-response. Ideally, the analysis would reflect a complete evaluation of all 5199 members of the defined cohort. However, loss of information occurs for several reasons, particularly in a study such as this one where the exposure occurred so long ago: inability to locate people, refusal to participate, inability to participate for other reasons, and mortality within the cohort. The primary concern with such losses is the possibility that people who are not included in the final analyses are somehow different in a systematic way that might be related to both: 1) radiation dose from Hanford; and 2) one of the thyroid outcomes under study. If so, failure to include such people could potentially result in a misleading or incorrect estimate of any dose-response relationship.

This study was successful in locating members of the cohort. As reported in section V.B.4., approximately 94% of the 5199 individuals originally identified were located: 4350 alive and 527 deceased. Furthermore, there was no evidence of an appreciable difference in ability to locate people according to sex, year of birth (and thus age at exposure), or geographic area of birth within the Hanford region. The proportion located in each of these subgroups was high (over 90%), and relatively uniform.

Once located and contacted approximately 84% of those contacted agreed to participate. This high level of cooperation was also relatively uniform among the various subgroups defined by sex, year of birth, and geographic area of birth. There was no evidence of any subgroup of the cohort being substantially more or less likely to agree to participate. This pattern also was apparent according to geographic area of current residence. There was no indication that people who live outside of the region were less likely to agree to participate. Across all areas of the country, the proportion agreeing to participate was uniformly around 80%.

Approximately 15% of those contacted refused participation, or withdrew from participation (even though they initially agreed to participate). An attempt was made to identify a reason for each refusal or withdrawal based on responses to the refusal questionnaire (if the person was willing to provide such information) and the recruiter's assessment of the interaction with the person. The majority of the refusals

and withdrawals were because the person was not interested, or did not have time. Very few refused because they were opposed to the study. Potentially of more interest is the group who refused or withdrew because of illness or impairment, which precluded them from participating. The principal concern would be that people in this group were more likely to have one of the outcomes under study. In only one case was current thyroid disease given as a reason for non-participation. Further, in reviewing the types of illnesses and impairments cited, there is no indication that people who refused or withdrew were more or less likely to have any of the disease outcomes investigated in the study. Although it is not possible to say with certainty from these data that people who chose not to participate are no different regarding the outcomes of interest than those who did participate, these responses provide some assurance that such is the case. In addition, given the relatively small proportion of people who refused or withdrew, and the uniformity of this proportion among subgroups of the cohort, it is unlikely that such losses could have materially biased the dose-response analyses.

Loss of information also occurs because of mortality within the cohort. Of potential concern is the possibility that such loss is related to one or more of the outcomes of interest. In this study, 527 (10.1%) of the 5199 individuals originally identified were confirmed as deceased and an additional 16 (0.3%) who were located alive died before participating in the HTDS. A death certificate was obtained for 504 (93% of the 543) in order to determine the cause of death for each person. There were 199 deaths in females and 344 deaths in males, with no known age of death for two of the males. For both sexes, the largest proportion of deaths occurred under one year of age (36% for both males and females). Most of these deaths were due to conditions in the perinatal period or congenital anomalies. Approximately 31% of the deaths in females were due to these two causes, as were approximately 27% of the deaths in males.

An analysis was conducted to investigate whether the mortality experience in this cohort overall was unusually high, relative to what would be expected based on the mortality experience of the regional population from the same time period, and to determine whether there was any indication of an excess in mortality from conditions that might be related to one or more of the primary thyroid outcomes of interest. The detailed results of this analysis are shown in Mortality Appendix 23. In summary, there was no overall increase in total mortality over what would be expected based on the mortality experience of the population of Washington State during the same time period (standardized mortality ratio (SMR) = 0.97; 95% Confidence Interval (CI) = 0.89, 1.06). This was true for both females (SMR = 0.96) and males (SMR = 0.98). However, there was an excess in deaths due to conditions of the perinatal period (SMR = 1.69, 95% CI = 1.39, 2.04), which was found in both females (SMR = 1.70) and males (SMR = 1.68).

Further analyses were performed to investigate whether there was any excess in mortality according to geostratum of birth, or in birth years concentrated around the time of the peak releases from Hanford (i.e., in the birth cohorts defined by the period 1945-46). The only excess in mortality observed by geostratum was among people born in Franklin County (SMR = 1.61, 95% CI = 1.15, 2.20). This excess was found for males (SMR = 1.66, 95% CI = 1.09, 2.44), but was only suggestive for females (SMR = 1.53, 95% CI = 0.83, 2.56). There was essentially no difference in mortality seen between the 1945-46 birth cohorts and the 1940-1944 birth cohorts. Analyses were also conducted according to year of death, classified as before 1945 (beginning of Hanford operations) and 1945 or later. For total mortality, there was little difference in the SMRs for deaths before 1945 and for the period from 1945 on (SMR = 1.06 vs. 0.95, respectively), and neither was statistically significant. This pattern was similar in males and females, and was observed for conditions of the perinatal period and for congenital anomalies.

Of primary interest in considering the results of these exploratory analyses regarding mortality in the HTDS cohort is whether the loss of cohort members through death could in some way bias the dose-response analyses. Given the principal findings of the study, the primary concern would be that this loss attenuated a true dose-response (i.e., biased the estimate of effect toward the null), and that is why no association is observed between increasing radiation dose from Hanford and the outcomes under study. In order for the exclusion of participants lost to death to mask a true dose-response, one of three circumstances would have to be operative among the group of 543 deceased individuals: 1) they would have had to have experienced disproportionately higher doses *and* higher rates of the outcomes under study, thereby "pulling up" the high end of the dose-response curve; 2) they would have had to have experienced disproportionately low doses *and* low rates of the outcomes under study, thereby "pulling

down" the low end of the dose-response curve; or 3) they would have had to have experienced the full range of possible doses, but exhibited a *very strong* dose-response over the full dose range. Given the relatively small proportion of the cohort lost due to deaths (10%) and the consistency of findings of an absence of a radiation dose-response across all outcomes, it is highly unlikely that the absence of a dose-response is due to any of the these three circumstances. In fact, there is little difference in overall mortality among cohort members compared to what might be expected based on mortality rates in the same region over the same period of follow-up, and no evidence to suggest that cohort members born in the years of the peak radiation releases from Hanford experienced higher than expected overall mortality.

An alternative but related explanation of how the loss of deceased members of the cohort could attenuate a dose-response might be that those who died were somehow more likely to have developed thyroid disease had they lived (sometimes referred to as a "healthy survivor" effect), or perhaps died with undiagnosed or unrecognized thyroid disease. However, in order for such explanations to contribute to or account for the absence of a dose-response, it must be assumed that those who do not survive experienced high or at least appreciable doses. Although it proved impractical to estimate individual doses for the deceased in this study, a number of additional analyses were undertaken among subgroups of people defined by cause of death, year of death, area of birth, and time of birth in an attempt to investigate patterns of mortality that might conceivably be related to thyroid dose. Mortality in excess of that expected occurred primarily at very young ages, concentrated in causes related to the perinatal period, but the excess was apparent for deaths that occurred prior to the beginning of Hanford operations, and were similar in magnitude to the excess seen for deaths that occurred in 1945 or later. This would argue against the cause of such excess at young ages (or conditions of the perinatal period) to be related to exposures from Hanford operations. Similarly there was little evidence that any excess mortality was concentrated in people more likely to experience higher dose based on geography. Finally, there was no indication that deaths were concentrated in categories that might be related to the development of the outcomes under study, and there was no mention of thyroid diseases on any of the death certificates. Thus, although it is not possible to know whether those who died had higher doses and might have been more likely to develop thyroid diseases had they lived, or had unrecognized thyroid disease at the time of their death, there is no indication of such based on examination of the data available.

In summary, a total of 3440 (66.2%) of the 5199 individuals initially selected for inclusion in the HTDS cohort were evaluable and provided data for the analysis. The proportion of those originally selected who attended a clinic and were evaluable was remarkably uniform across the factors that defined the selection: sex (males 64.1%, females 68.3%), year of birth (1940: 67.3%, 1941: 69.5%, 1942: 69.3%, 1943: 66.6%, 1944: 65.3%, 1945: 63.0%, 1946: 66.2%), and geostratum (Richland: 64.9%, Pasco/Kennewick: 64.8%, Walla Walla City: 64.1%, Benton County: 65.2%, Franklin County: 63.7%, Walla Walla County: 71.7%, Okanogan: 65.9%, Ferry/Stevens: 63.9%, Adams: 73.8%). Thus, although the final dose-response results are based on approximately two-thirds of the people originally identified for study, it appears that the degree of loss of individuals from the group was relatively uniform across subgroups defined by sex, year of birth, geographic area of birth, and geographic area of current residence. There is no indication that people were less likely to participate because they had thyroid disease, and in more general terms, illness was infrequently given as a reason for non-participation. Further, there is no indication of a substantial loss due to mortality in ways that are likely related to both exposure (dose) and the development of any of the outcomes of interest (i.e., in ways that would substantially affect the estimates of dose-response). Although one cannot rule out the possibility that the dose-response results might be biased in some way as a result of non-participation by nearly one-third of the cohort, no patterns of non-response or loss to follow-up are apparent from the data available that would suggest such is the case. In order for such a bias to have an important influence in producing the pattern of results seen in the HTDS (lack of a dose-response), one would have to postulate that people who did not participate were more likely to have one of the outcomes under study *and* to have received higher doses. As noted above, there is no evidence of such selection bias in the HTDS cohort.

C.2. Factors Related to Outcome Definition

An important element of a cohort study such as the HTDS is that the outcomes of interest are ascertained in a comprehensive and unbiased manner. That is, it is important that all cases of a given outcome are identified, and that the identification of cases is not influenced by or related to exposure or dose. The clinical component of the HTDS was designed to ensure that such was the case. Because of the long time period between the Hanford exposures and present day, and because the thyroid diseases under study can often be difficult to diagnose or even go undetected, it was felt essential that each participant undergo a thorough examination and evaluation for the presence of each of the outcomes under investigation as part of their participation in the study. Great care was taken to ensure that each person received the most complete evaluation possible by using highly experienced thyroid specialists. Further, two different physicians examined each participant separately, consulted with each other, and reached agreement on their findings before the participant left the clinic. State of the art technology was used in the form of thyroid ultrasound to help ensure that all thyroid nodules were identified. Nearly all participants who attended a study clinic completed all aspects of the evaluation, including providing a blood sample for laboratory tests and undergoing a fine needle aspiration biopsy when recommended. Analyses of the "pathways to diagnoses" of thyroid cancer, benign thyroid nodules, and nodules suspicious for follicular neoplasm demonstrated that the numbers of such cases were increased by the comprehensive clinical evaluation provided to each participant. Given this design, and the success experienced in carrying out the clinical component of the study, it is felt that the ascertainment of outcomes in the cohort is essentially complete. It is highly unlikely that substantial numbers of cases of any of the primary outcomes of interest were undetected, or that there is any substantial misclassification of outcomes.

The study was also designed to try to minimize the possibility that the physicians or sonographers could be influenced in their evaluation by knowledge of the participant's possible level of exposure to Hanford radiation. As outlined in section V.F.2.d, a number of measures were taken to prevent this from happening. At the clinic, participants were instructed not to make the physicians or sonographer aware of any personal circumstances that would suggest what their radiation exposure history might be. They were also asked not to wear clothing items that might provide any such indication. A variety of clinic locations were used, and participants were scheduled into clinics in a way that purposely *did not* correspond to prior residence history or the likelihood of exposure. Thus, when an individual physician examined a participant, he had no knowledge of what that particular participant's past history was in relation to the Hanford radiation releases (or any other potential radiation exposures). The same was true for the sonographer. As a check to see whether these precautions were effective in blinding the physicians to possible exposure, each physician was required to indicate at the conclusion of their evaluation whether they had any indication of possible exposure for that individual. In only 15 instances (of the 3440 living evaluable participants) did the physician suspect some knowledge of past exposure. Precautions were also taken to blind the physician reviewers of past medical records to any mention of radiation exposure. As described elsewhere, this was done in a manner that made it impossible for the reviewer to know for any given medical record whether there was any indication of previous exposure to radiation (either from medical or environmental sources). In summary, based on the success of the various approaches used, it is not likely that the determination of outcomes was influenced in any substantive way by knowledge of exposure.

C.3. Factors Related to the Estimation of Thyroid Radiation Dose

Just as it is important to accurately define outcomes, it is critical to accurately classify study participants according to exposure or dose. In this study, substantial misclassification of study participants according to radiation dose would tend to attenuate any true dose-response relationship (i.e., bias the estimate of effect towards the null). One approach to minimizing the likelihood of substantial exposure misclassification in a study such as this one where dose is estimated (reconstructed) based on historical information is to utilize individual-level information as much as possible to "tailor" each individual's estimate of dose to his or her own specific circumstances. The HTDS was designed from the beginning to use this approach. The cornerstone of the method was to elicit detailed information for each respondent regarding those factors most crucial in determining thyroid radiation dose from Hanford, and to use that

information to estimate an individual dose for each person. This was difficult to do because of the level of detail required, the long period of time that has elapsed since the exposures, and the fact that participants were young children during the time period that is most relevant for estimating dose. A considerable effort was made to structure the collection of individual information in a way that would enhance a person's ability to recall the information accurately by using a cognitive approach to interviewing, and to do so in a manner that would not be biased by the participant's knowledge of thyroid disease status. The administration of both the CATI and In-Person Interviews prior to the clinical evaluations probably aided in avoiding bias in recall to some extent. However, there is no way to directly assess the degree of potential misclassification of exposure that occurred using the approaches taken in this study.

Therefore, we repeated all of the analyses using alternative methods of assigning exposure to see if the results changed in any substantial way. First, we developed two alternative dose schemes that maintained an individual level dose estimate for each participant, using the HEDR models to estimate dose. The primary analyses were based on doses estimated using individual residence histories, individual responses to the CATI (or Expanded In-Person Interview), and HEDR default values when CATI responses were not available. The first alternative individual dose scheme used individual residence histories, and HEDR defaults exclusively (that is, no data from CATI or Expanded In-Person Interviews). The second alternative dose scheme was the same as the primary scheme, but HTDS default values were used when CATI responses were not available instead of HEDR default values (with the exception of consumption other than milk for Expanded In-Person Interviews, for which HEDR defaults were used). These HTDS default values for food and milk consumption data were defined based on the experience of the participants in the HTDS for whom a completed CATI interview was available. None of the dose-response results for any of the outcomes changed appreciably from the primary results using either of these alternative methods of estimating individual thyroid dose. This provides some assurance that the absence of a dose-response found in the primary analyses is not due to misclassification of exposure introduced by difficulties in recall from the distant past.

Second, two alternative representations of exposure were defined which were independent from the HEDR dosimetry system altogether, and therefore did not use the HEDR models to estimate individual radiation dose. One was simply the geostratum used to define the sampling frame for selecting the cohort (i.e., the mother's usual place of residence at the time of the participant's birth as determined from the participant's birth certificate). Although this is clearly an imperfect surrogate indicator for Hanford radiation dose, and does not take into account individual circumstances (e.g., movement patterns and dietary habits), it might provide at least a crude way to distinguish people more or less likely to have received substantial exposures.

For the primary definition of each outcome, analyses were conducted to see whether there was heterogeneity of outcomes across geostrata, and whether the proportion with the outcome in the two geostrata defined by Okanogan and Ferry/Stevens counties was different than that in the remaining seven geostrata. In summary, there was little evidence of significant heterogeneity in the cumulative incidence of any of the outcomes across all geostrata. Those outcomes showing the greatest degree of variation across geostrata were benign nodules, any thyroid nodules, any thyroid UDAs, and palpable thyroid UDAs.

Of more interest was the generally consistent finding that the proportion of participants with a given outcome was somewhat higher in the Okanogan and Ferry/Stevens geostrata than in the other seven geostrata. This pattern was apparent for the most part across all primary outcomes, although for those with very few cases (e.g., thyroid cancer, simple goiter, Graves disease, hyperparathyroidism) there was very little statistical power to evaluate the relationship. Insofar as geostratum serves as a surrogate indication of radiation exposure (and dose), and the underlying hypothesis is that radiation exposure from Hanford is associated with an increase in the thyroid disease outcomes under study, these results were quite unexpected because the Okanogan and Ferry/Stevens geostrata were defined in an attempt to identify people who were likely to have been relatively *unexposed* to Hanford radiation releases. Indeed, according to the individual dose estimates derived using the HEDR models, it appears that the sampling strategy was successful in that regard because the average doses for living evaluable in area participants in these two geostrata were the lowest of all nine geostrata (see Table IX.B-4, section IX.B: Okanogan, 11 mGy; Ferry/Stevens, 36 mGy).

It is not readily apparent why the cumulative incidence of the thyroid diseases under study would be slightly higher among people born in these three counties. Based on individual dose estimates that account for a person's movements and lifestyle, it appears that those selected from the Okanogan and Ferry/Stevens geostrata have the *lowest* doses in the cohort. There is also no evidence that this group is unusual in terms of selection or participation in the study, or ascertainment of disease status. Further, all of the analyses by geostrata were adjusted for differences by sex and age at examination. It is also difficult to imagine that some other aspect of birth, early life, or living in these areas is related to the risk of developing thyroid disease, as the apparent effect is seen across *all* outcomes (including hyperparathyroidism and thyroid UDAs). One would have to postulate that such an influence is related to all the different forms of thyroid disease included in this study, which seems exceedingly unlikely. A possible exception might be iodine deficiency. Geographical differences in the distribution of iodine intake (e.g., endemic goiter belts) could result in geographic differences in the rates of one or more of the thyroid diseases under study.

As described in section IV above, there is very little information available describing either estimates of soil iodine concentrations or iodine intake on a geographical basis. Probably the most useful data in this regard are those reported by Oddie et al. (119). He reported estimates of average dietary iodine intake derived from thyroidal radioiodine uptakes in approximately 30,000 euthyroid subjects in 133 locations throughout the United States. Although average daily iodine intake varied considerably throughout the United States (from 240 to 740 micrograms per day), the Pacific Northwest was relatively uniform in the distribution of daily intake estimates. Mean values were reported for fifteen areas in the Northwest centered by two degrees latitude and longitude (approximately 140 by 120 miles). All values in the six HTDS Pilot Study counties were between 345 and 379 micrograms per day (a very narrow range compared to the overall distribution of values). These findings provide some evidence that iodine intake was adequate and relatively uniform in the past in the areas from which study participants were selected. As such, they suggest that iodine deficiency is not a likely explanation of the relatively higher proportions of thyroid disease among people selected from the Okanogan and Ferry/Stevens geostrata.

Nevertheless, because the cumulative incidence of disease was consistently higher in the Okanogan and Ferry/Stevens geostrata in a manner possibly related to dose, it was decided to repeat the primary dose-response analyses omitting people born in these two geostrata. If thyroid disease rates were truly elevated in the population from which people in these geostrata were selected, and such people tended to have lower Hanford doses, the dose-response analyses might be biased toward the null (i.e., the dose-response might be underestimated). Generally, the effect was to increase the regression coefficient (slope of the dose-response). However, none of the changes were substantial enough to suggest a significant dose-response relationship. The largest changes were for the outcomes related to thyroid UDAs. For any thyroid UDAs, the regression coefficient increased from 0.031 per Gy to 0.046 per Gy, and the p-value changed from 0.21 to 0.11. Thus, although the effect of excluding participants from these geostrata had the anticipated effect on the dose-response results, it did not materially change the overall findings or conclusions. It should be emphasized that analyses that excluded the Okanogan and Ferry/Stevens geostrata were not included in the original analysis plan, but were conducted only after the higher cumulative incidence rates in these two geostrata were observed. Given the data-driven nature of this additional analysis, there is no evidence to suggest that the somewhat higher cumulative incidence of disease in the Okanogan and Ferry/Stevens geostrata led to a significant underestimate of the dose-response for any of the primary outcomes under study.

The second alternative representation of exposure which did not use the HEDR models to estimate individual dose was based on the assumption that two factors are particularly important in determining radiation dose from Hanford: a person's residence history and history of milk consumption. A dichotomous representation of possible exposure (high, low) was defined based on this information. For this analysis, the high exposure group was defined to include those living evaluable participants who:1) were born prior to July 2, 1945;2) lived for at least 180 days in Benton, Franklin or Adams counties (excluding Richland) during 1945; and 3) consumed on average at least one serving of milk per day during 1945. The low exposure group was defined to include: 1) all out-of-area participants (OOA); 2) participants who lived only in Ferry, Stevens or Okanogan counties or OOA in 1945 and who never lived in Benton, Franklin or Adams counties between 1946 and 1951 inclusive;3) participants born in 1946 who never lived in Benton, Franklin or Adams counties between 1946 and 1951 inclusive; or 4) participants who lived outside of

Benton, Franklin or Adams counties from the later of the participant's birthday and 12/15/44, until 12/31/51, and consumed on average less than one serving of milk per day in 1945 (includes only participants with CATI as the dose source).

Using this dichotomous representation of possible exposure (high, low) and the primary definition of each outcome, analyses were conducted to see whether the cumulative incidence of each outcome was greater among those in the high dose category relative to the low dose category. In summary, there was no evidence of a significant relationship between exposure, as represented in this manner, and the cumulative incidence of any of the outcomes. There was a slightly higher proportion of participants with thyroid UDAs in the high group relative to the low group (50.3% vs. 47.4%), but not significantly so ($p = 0.10$), and this relationship was somewhat more pronounced when the analysis was restricted to nonpalpable focal thyroid UDAs (41.5% vs. 37.6%, $p = 0.079$).

Although this approach for assigning exposure is also crude, it incorporates at least some of the information about each individual's circumstances that is thought to be important in the determination of dose. As such, this surrogate indicator should be somewhat more capable of distinguishing people who received relatively higher doses from those who received relatively lower doses than the simple geostratum designation. If so, it nevertheless does not provide any evidence of a statistically significant association between higher Hanford radiation dose to the thyroid and an increase in any of the primary outcomes under study.

A limitation of the dosimetry system available for this study was its inability to calculate dose estimates for participants who did not live within the HEDR domain between December 1944 and the end of 1957. As a result, the primary dose-response results of this study refer to dose received while living in the HEDR domain between December 1944 and the end of 1957. Individual dose estimates could not be calculated for the 249 participants who lived outside the HEDR domain during that period, the so-called "out-of-area" participants. It is reasonable to assume that the out-of-area participants received generally low doses. In particular those who lived only at great distance from Hanford during this time period probably received virtually no dose from Hanford. However many out-of-area participants lived in places not far outside the HEDR domain. It is probably inappropriate to simply assume that such people received no exposure from Hanford.

Therefore, scoping analyses were performed to assess whether inclusion of the out-of-area participants in the primary analyses, had that been possible, might have substantially changed the dose-response results. These analyses assigned crude estimates of a dose for the out-of-area participants, based on residence during the 1944-1957 exposure period. Out-of-area participants who lived in the four states or two Canadian provinces closest to Hanford were assigned doses of either 0 mGy, or the highest dose that they would have been assigned had they lived on the border of the HEDR region in the direction of the state or province (which would likely overestimate the dose they could have actually received), depending on their disease outcome status. Those who lived outside that four-state/two-province region were assigned doses of 0 mGy. A scoping analysis of each disease outcome was then performed in which all out-of-area participants with the outcome were assigned their "border dose," while those without the outcome were assigned 0 mGy. This imposes a strong dose-response relationship among the out-of-area participants. However when the in-area and out-of-area participants were combined in these scoping analyses, there were no important changes in the estimated dose-responses. This was true even in the analysis of thyroid cancer, for which five of the 19 cases based on HTDS or prior histology were out-of-area participants. A second scoping analysis assigned doses in the reverse order, so that out-of-area participants with the outcome received a dose of 0 mGy and those without the outcome their "border dose." This did not materially change the estimated dose-response for any outcome either.

It is perhaps not surprising that neither of the scoping analyses which included the out-of-area participants had much impact on estimated dose-responses, since the out-of-area participants comprised only 7.2% (249/3440) of the living evaluable participants. Moreover the crude dose estimates that they were assigned ranged from 9 to 48 mGy, well below the mean dose of 174 mGy observed for the 3191 in-area participants.

In summary, a number of attempts were made to use alternative approaches for characterizing study participants in terms of their exposure to Hanford radiation, including both alternative quantitative and qualitative schemes. This was done so that the investigation of a possible relationship between Hanford radiation exposure and thyroid disease would be as complete and comprehensive as possible, would rely on multiple types and sources of data, and would not be limited to only one dose assessment approach and the associated assumptions. It was recognized from the beginning of the study that there would be limitations in the quantitative dose estimation program developed by HEDR, and that alternatives based only on residence location would provide crude indicators of exposure at best. The decision to use both approaches, and to look for consistency in results, was felt to provide a more thorough assessment of a possible relationship between radiation exposure and thyroid disease. Analyses of all of the primary outcomes were repeated for each alternative approach. None of these analyses produced evidence of a statistically significant relationship between any of the primary outcomes and exposure to Hanford radiation (or dose). The principal findings of the primary analyses using individual doses estimated by the HEDR dosimetry system were not materially changed by any of these alternative analyses. In addition, all primary analyses were repeated using the arithmetic mean, and the geometric mean, of the 100 dose realizations for each participant rather than the median dose estimate. This did not change the results. All primary analyses were also conducted fitting sex-stratified linear-quadratic and logistic dose-response models. The addition of the quadratic term did not significantly improve the fit of the dose-response model for any of the outcomes under study, and neither the linear-quadratic or logistic models provided any evidence of a significant radiation dose-response.

For many of the disease outcomes, the numbers of cases among participants with the highest dose estimates tended to be relatively low. As a result, estimated slopes of the dose-response relationships were slightly, though not significantly, negative for these outcomes. Additional analyses were performed to assess whether these results might be unduly influenced by the relatively small proportion of participants with the highest doses. In particular, the primary analyses of disease outcomes were replicated twice: once excluding participants with estimated doses above 1000 mGy, and a second time excluding those with doses above 400 mGy. The first alternative analysis had very little impact on the fitted dose-response models. A somewhat stronger impact was seen in the analysis that excluded participants with estimated doses over 400 mGy. For most disease outcomes, the slope of the dose-response tended to be greater when based on the limited set of participants, although in general the increases were not large enough to suggest a statistically significant dose-response. For two outcomes, the exclusion of participants with estimated doses over 400 mGy increased the estimated slope of the dose-response substantially. For nonpalpable focal thyroid UDAs the estimated slope increased from 0.027 per Gy ($p = 0.23$) to 0.228 per Gy ($p = 0.003$). For diffuse thyroid UDAs the estimated slope increased from 0.029 per Gy ($p=0.14$) to 0.146 per Gy ($p=0.005$).

While the magnitudes of the dose-responses for these two ultrasound outcomes excluding participants with estimated doses over 400 mGy are considerably larger than the estimates among all study participants, the statistical significance of these results must be interpreted with caution. First, this is a secondary, exploratory analysis that only shows a significant effect when people with the *highest* thyroid doses are excluded. Second, it should be kept in mind that this result was found in the context of conducting many secondary and alternative analyses and significance tests. Third, such abnormalities are quite common. Numerous investigations in populations throughout the world have reported that 20-50% of individuals may have one or more such findings on ultrasound examination. Fourth, and perhaps most importantly, the health significance of nonpalpable focal and diffuse thyroid UDAs is unclear. Whereas thyroid UDAs that are palpable can be classified as thyroid disease, the high prevalence of those that are not palpable may not represent clinical disease. Since no dose effect was detected for recognized thyroid disorders such as thyroid cancer, benign thyroid neoplasia, and hypothyroidism, it would seem unlikely that the focal and diffuse ultrasound findings would be clinically significant. Could these ultrasound findings represent *subclinical* thyroid disease? In other words, very mild abnormalities that do not cause symptoms but might be destined to become clinical disease over time? If this were true, one might expect to see 2 types of dose-response results in the HTDS: an increase in the number of ultrasound abnormalities with increasing dose, and an increase in the risk of having an ultrasound abnormality of a particular size with increasing dose. The HTDS examined both of these possibilities. First, there was no relationship found between the number of ultrasound abnormalities on a participant's ultrasound scan with increasing dose.

Second, there was no increased risk of having larger, focal ultrasound abnormalities (maximum size 5mm, maximum size 10 mm, or average size at least 15 mm) with increasing dose. Thus, these results do not suggest that these ultrasound findings represent early manifestations of thyroid disease. In summary, based on the above factors, it would seem very unlikely that the dose-response seen for nonpalpable focal and diffuse ultrasound abnormalities, found only after secondary, exploratory analyses, and only after excluding participants with the highest doses, truly represents a significant dose effect in the HTDS.

C.4. Potential for Confounding or Effect Modification

Although relatively few factors have been well established as important in the etiology of the thyroid diseases under study, an attempt was made to collect as much information as possible from study participants regarding aspects of their personal history and lifestyle that might potentially influence the risk of developing thyroid disease. As described in section VIII, this information was used to construct several variables for inclusion as covariates in the dose-response analyses. Analyses were conducted for all outcomes with sufficient numbers of cases to evaluate whether any of these factors confounded the relationship between the outcome of interest and estimated Hanford radiation dose, or whether any dose-effect was modified by levels of the factor (e.g., for sex, whether the effect was different in males and females).

None of the covariates investigated materially changed the estimates of the dose-response for any of the outcomes under study. There was no evidence of confounding by any of the factors, nor was there any evidence of effect modification by any of the factors assessed. This included the covariate reflecting exposure to radiation from the Nevada Test Site. These rather extensive analyses provide no evidence that there is a significant dose-response for any of the outcomes under study, or evidence of a significant dose-response among subgroups of participants defined by any of the covariates investigated.

C.5. Statistical Power of the Study

Of critical importance in the interpretation of these results is the ability of the study to detect an increase in disease risk if it is present, i.e., the statistical power of the study. In order for the findings of an absence of an effect to be very meaningful, there must be adequate statistical power to detect an effect of the magnitude that might be expected based on existing knowledge, and that is relevant and meaningful to the population exposed. As described more fully in section VIII, the HTDS was designed to have relatively high power to detect a positive dose-response as small or smaller in magnitude than any existing published findings regarding each outcome. These projections of study power, which were based on the results of the Pilot Study, were actually exceeded in the Full Study (as shown in Table IX.B-14 above). Nevertheless, because uncertainties in the individual dose estimates could be expected to reduce study power, we undertook a simulation analysis to estimate the impact on study power of incorporating such uncertainties in the dose estimates (see section IX.B-4). Although the effect of dose uncertainty was, as expected, to reduce the statistical power of the study, the reduction was modest. Even after accounting for uncertainty in doses, the HTDS had greater than 80% power to evaluate each of the hypotheses originally specified.

To interpret the study's power properly, it is important to consider not only the level of power, but also the size of the dose-response effect for which that power is obtained. As described in section IX.B.4 above, after accounting for the impact of dose uncertainty, the study's one-sided tests at critical level $\alpha = 0.05$ had estimated power of about 85% to 86% to detect linear dose responses corresponding to relative risks (average for both sexes) of 2.04, 1.30, and 1.05 at the study participants' average dose of 174 mGy, for the exemplary outcomes with low (thyroid cancer), intermediate (any thyroid nodule), or high (thyroid UDA) background rates, respectively.

For comparison to results of other studies, the magnitudes of radiation effects can be expressed as the relative risks at 1000 mGy (1 Gy). For the low background rate example of thyroid cancer, a slope of 2.5% per Gy, for which HTDS had about 86% power (Table IX.B-16above), corresponds to a relative risk

(average of both sexes) of 6.95 at 1 Gy. This is a substantially smaller effect than that observed in the Utah Thyroid Study, for which the relative risk was estimated as about 25 at 1 Gy after accounting for dose uncertainties (134). A recent analysis suggested that the adjustment should perhaps be smaller: Mallick and colleagues analyzed the Utah Study's data concerning thyroid neoplasms and concluded that the estimated relative risk at 1 Gy should be approximately doubled, rather than tripled, to account for dose uncertainties (140). Assuming this conclusion applies to thyroid cancer, the estimated relative risk would be about 17 at 1 Gy. The HTDS clearly had adequate statistical power to detect an effect of this magnitude. For example, after accounting for dose uncertainty there was an estimated 92% power to detect a linear dose-response with a slope of 3.5% per Gy for thyroid cancer (Table IX.B-16 above), which corresponds to an average relative risk (both sexes combined) of 9.33 at 1 Gy, well below the estimated effect from the Utah Study.

D. Comparison of Results with Findings in Other Populations Exposed to Radiation

Although there is a substantial literature regarding the role of ionizing radiation in the induction of thyroid disease in humans, the findings reported to date do not provide a clear and consistent characterization of the relationship between radiation exposure and risk. This is due in part to the fact that a number of factors are probably important in determining risk: the type of radiation, the dose received, the rate at which the dose was received, a person's age at the time of exposure, a person's age at the time of disease occurrence, and iodine deficiency. Thus, in comparing the results of the present study with those published, it is important to keep in mind the characteristics of the Hanford exposures and the basic design features of the HTDS. The exposure was environmental, and occurred over a period of up to approximately 13 years, although much of the dose was delivered in a considerably shorter period of time, and many people may have received most of their dose over periods of several months. The design of the HTDS resulted in a study group that consisted of people who were young children (under age 5) at the time of the peak exposures, and follow-up occurred over a period of up to more than 50 years. Radiation dose to the thyroid from ^{131}I was estimated for each individual, based on historical reconstruction of events. Estimated doses for the study group were relatively low (median dose = 97 mGy, mean dose = 174 mGy). Thus, it is within this context that the present results are considered in relation to the published literature. The primary goal of this comparison is to evaluate how well the current findings "fit in" with what is currently known about radiation-induced thyroid disease. To the extent possible, specific analyses have been tailored to be as comparable as possible to published results, for the explicit purpose of direct comparison.

A more detailed presentation of the published literature is contained in sections II.B through II.D above. It is not the intent to repeat those descriptions here, but rather to highlight the principal points for comparison with the HTDS findings. There is clear evidence from a number of studies that people exposed to external sources of gamma radiation or x-rays are at an increased risk of developing thyroid neoplasia. There is also evidence to suggest that the risk is greater for people exposed at younger ages. Most of this evidence comes from studies of people treated medically with radiation, and from studies of the survivors of the atomic bombings in Japan. Thus, in both circumstances, doses were generally considerably higher than those in the HTDS, were generally delivered at a much higher dose rate, and reflect external exposures. Nevertheless, one study of children irradiated for *tinea capitis* provides some evidence of an increased risk associated with much lower doses (average dose = 90 mGy).

Of much more relevance to the Hanford circumstances are studies which have evaluated the effects of exposure to radioactive iodine. Unfortunately, much less information is available in this regard, especially in human populations. Two types of information exist: findings based on people exposed in medical settings, and findings based on people exposed environmentally. People exposed therapeutically to radioactive iodine (primarily for the treatment of Graves disease) generally received very high doses. However, there is no clear evidence that such exposures result in a subsequent increase in thyroid neoplasia. People exposed for diagnostic purposes generally received much lower doses, but the doses are still relatively high compared to the Hanford doses (typically 500 – 1000 mGy). There is no convincing evidence that exposures at these levels result in increased thyroid neoplasia. Although the rates of thyroid cancer were elevated in some of the above studies, the authors concluded that the increase was more likely related to the underlying thyroid disease than to the radioiodine exposure.

Information regarding the effect of environmental exposure to radioactive iodine comes from studies of three principal populations: people exposed to fallout from atmospheric nuclear testing in the Marshall Islands in the 1950s, people exposed to releases from the Chernobyl Power Station accident in the Former Soviet Union in 1986, and people exposed to fallout from atmospheric nuclear testing at the Nevada Test Site in the 1950s and early 1960s. The experiences in the Marshall Islands and at Chernobyl are less directly comparable to the Hanford experience because the exposure in each instance consisted of a broader and different mixture of radionuclides, and the dose rates were relatively high (short time of exposure). Nevertheless, in the Marshall Islands there has been an increase observed in thyroid neoplasia associated with the more highly exposed areas, with doses much higher than those around Hanford. Around Chernobyl there has been reported a dramatically increased occurrence of thyroid cancer in young children. Unfortunately, there are no epidemiologic studies available with quantitative estimates of individual thyroid radiation dose from Chernobyl to better elucidate the nature of any dose-response in this regard. However, a number of attempts to estimate radiation doses on a population basis suggest that the doses were generally much higher than those around Hanford.

The study of people exposed to fallout from the Nevada Test Site, the so-called "Utah Study"(133, 134), is probably the most comparable to the Hanford situation. The mean dose for all 3545 participants who were included in any phase of the Utah Study was 98 mGy, compared to 174 mGy for the 3191 living evaluable in-area HTDS participants. The maximum estimated thyroid dose in the Utah Study was 4600 mGy (2823 mGy for HTDS), although only 10 participants (0.3%) had estimated doses greater than 1000 mGy (24 or 0.8% for HTDS). However, there was likely a greater contribution from short-lived radioiodines and external radiation in the Nevada Test Site exposures compared to exposures at Hanford. Moreover, the participants in the Utah Study received most of their dose in short time periods after one or more test detonations. In contrast, most Hanford exposures were continuous and prolonged over months or years. A statistically significant dose-response was reported for total neoplasms (benign follicular neoplasms and thyroid cancer) in the 2473 participants who were included in the Utah Study's analysis of period prevalence between 1965 and 1986. Based on the linear relative risk model, the excess relative risk was estimated to be 0.070 per mGy, with unadjusted 95% confidence interval ranging from 0.007, 0.33 per mGy ($p = 0.019$). A relative risk of 3.4 (95% confidence interval 0.5, 26.9) was reported for all thyroid neoplasms for people with a dose of greater than 400 mGy. Although there were positive dose-responses for thyroid cancer and total nodules when these two outcomes were analyzed separately in the Utah Study, they were not statistically significant ($p = 0.16$ and 0.096, respectively).

Analyses that adjusted for the effect of dose uncertainties were also performed for the Utah Study. The dosimetry model and the approach to estimating doses for the Utah Study were, broadly speaking, similar to the HEDR model and HTDS approach. The size of the dose uncertainties was summarized as follows for the Utah Study: the geometric standard deviations (GSDs) for over 90% of the Utah Study participants were between 1.75 and 3.75 (133). This is generally similar to the magnitude of dose uncertainties for HTDS participants, whose GSDs ranged from 1.56 to 5.42, with a mean of 2.18 (see Section IX.B-2). The Utah Study investigators performed additional analyses in an attempt to adjust for the effect of dose uncertainties, which yielded adjusted estimates of the dose-response coefficients that were roughly three times greater than the unadjusted estimates. The standard errors of the estimates also increased in approximate proportion to the estimates, so the statistical significance of the dose-responses was essentially unchanged (133, 134). A recent reanalysis that attempted to account for the correlation of uncertainties in the Utah Study's dose estimates suggested that the adjustment should in fact have been somewhat smaller (139).

A number of other thyroid diseases investigated in the HTDS have also been linked to radiation exposure. It is clear that exposure to external gamma radiation, x-rays, or ^{131}I at high doses increases the risk of developing hypothyroidism. There is no evidence, however, that exposure to radioactive iodine, at lower doses similar to those estimated in the HTDS cohort, has the same effect. The HTDS found no statistically significant evidence of such an effect. This is consistent with the results of the Utah Study, which found no evidence that the risk of hypothyroidism increased with increasing estimated dose from the Nevada Test Site's fallout (10479).

Two recent studies have suggested that autoimmune thyroiditis may be radiation-induced. These findings come from studies of the Japanese atomic bomb survivors and people exposed around Chernobyl. As indicated above, they reflect very different types of exposures than at Hanford: external sources of exposure, higher doses, and higher dose rates. Nevertheless, for comparison purposes, we conducted an additional dose-response analysis that would correspond more directly to the analysis reported by Nagataki et al. (15). For that analysis we defined autoimmune thyroiditis to include only those cases associated with non-iatrogenic permanent hypothyroidism (see section IX.H above). The results of this analysis provided no evidence of a significant dose-response (slope of the dose-response = 0.001 ± 0.015; p-value = 0.48). It should be noted, however, that there were 161 cases of autoimmune thyroiditis in the HTDS cohort according to this definition (cumulative incidence of 4.7%), which is considerably higher than reported by Nagataki et al.. They report 27 clinical cases (1.0%) and 38 subclinical cases (1.5%) in their group of 2587. Unfortunately, insufficient detail is provided in the published paper to discern exactly how their cases were defined. Thus, it may be that the results of our alternative analysis are not truly comparable to those of Nagataki et al., and the reason that the HTDS was not able to confirm their findings may be in part due to the use of different criteria for the diagnosis.

The outcomes of hyperthyroidism, thyroiditis, and goiter were also investigated in the Utah Study, and for none of these were statistically significant dose responses observed (133). While these findings, taken at face value, appear to be consistent with the results of the HTDS, it is important to recognize that the definitions and diagnostic criteria used for these outcomes differed somewhat between the two studies.

There is also reasonably clear evidence that exposure to head and neck irradiation in childhood increases the risk of developing hyperparathyroidism. However, this evidence is based on situations in which the exposure was due to external sources and the doses and dose rates were generally quite high. There has been no convincing evidence in humans regarding the effect of exposure to radioactive iodine. However, it is estimated that the radiation dose to the parathyroid glands is less than that of the dose to the thyroid from a given exposure to radioactive iodine. Thus, given the thyroid dose distribution in the HTDS, it would be expected that parathyroid doses to members of the HTDS cohort were very low.

Relatively little is known about whether ionizing radiation causes an increase in thyroid abnormalities detected by ultrasound prior to the development of clinical disease. Schneider reported that exposure to external radiation was associated with a high prevalence of thyroid UDAs (112). In 54 exposed individuals followed in his study, 87% (47/54) had abnormal ultrasound scans. In this cohort, radiation exposure was due to external sources. The authors concluded that 1) thyroid nodules continued to develop in radiation-exposed individuals many years after exposure and 2) although thyroid UDAs were quite common in the general population, they were more prevalent in radiation-exposed populations.

Other studies have also suggested that thyroid UDAs are more common in exposed populations. Antonelli compared ultrasound scans among 50 hospital workers with occupational radiation exposure (external radiation) in a hospital setting to 100 controls without such exposure (113). Thyroid UDAs were detected in 38% of the exposed people and only 13% of the controls. Similarly, Sugenoya and colleagues (114) compared 299 children who were exposed to Chernobyl radiation to 323 children who were unexposed. Although none of the children in either group had palpable abnormalities, 34 of the exposed (11.4%) had thyroid UDAs compared to 4 unexposed children (1.2%).

There are no published estimates of the risk of developing thyroid UDAs as defined by the HTDS in relation to exposure to radioactive iodine.

Thus, in considering the HTDS dose-response findings in the context of the literature on radiation-induced thyroid disease, it is important to keep in mind the principal differences between the Hanford exposures and those in other populations that have been studied. The Chernobyl exposures occurred in a relatively short period of time and were substantially greater. Doses in populations around Chernobyl studied to date have generally been higher, and dose rates were much higher than at Hanford. The mix of radionuclides released was also different from Hanford, and there is some evidence that iodine deficiency may be contributing to the excess in thyroid cancer observed thus far. The Marshall Island experience is somewhat similar to the Chernobyl experience, insofar as doses were generally much higher and dose rates

much higher than at Hanford. The mix of radionuclides was also more varied than at Hanford. The exposures in Utah from the Nevada Test Site were also due to a broader mix of radionuclides than at Hanford, although resulting doses were similar to those at Hanford. The dose rate for any given individual in the Utah study was also relatively high, compared to that at Hanford, even though the exposures occurred over many years. That is because the exposures resulted from individual nuclear tests, which delivered the radioactive contamination in discrete, short periods of time. In contrast, exposures at Hanford were relatively constant over time (although concentrated in the early years of operation). Considered in total, the differences summarized above may largely explain why no dose-effects were observed in the HTDS analyses.

E. Comparison of the Occurrence of Thyroid Disease Outcomes With Other Findings in the Literature

The section above considered the thyroid disease dose-response results of the HTDS in the context of reported findings in other populations exposed to ionizing radiation. It is also important to consider the findings of the clinical component of the HTDS (the determination of thyroid disease outcomes) in relation to what is known about the occurrence of thyroid disease in other populations around the world. That is, how does the magnitude of thyroid disease occurrence found in the HTDS cohort (the cumulative incidence) compare with the levels of thyroid disease observed in other populations? Of particular interest is whether the occurrence of thyroid disease in the HTDS cohort is *greater* than has generally been found in other populations not exposed to ^{131}I from Hanford. If so, this might be considered evidence of a possible effect of Hanford radiation exposure, even in the absence of any dose-response relationships.

This is an exceedingly difficult question to answer because, as noted previously, the magnitude of thyroid disease rates observed in any given population depends upon a number of different factors. First, the recognition and diagnosis of thyroid disease in a population depends to a large extent on how aggressively one looks for disease. Sometimes referred to as the "screening effect", a concerted effort to screen for a disease in a population, including the implementation of a comprehensive diagnostic protocol as part of a research study like was done in the case of the HTDS, will result in higher rates of disease than would be observed with normal medical care practices in the same population. This is particularly so for thyroid neoplasia, which may not result in clinical symptoms and therefore can remain undetected, or functional forms of thyroid disease such as hypothyroidism which may go unrecognized as thyroid disease because of non-specific symptoms. Second, the extent to which thyroid disease is identified in a population may depend on the diagnostic methods used or the criteria for diagnosis that are employed. For example, the use of thyroid ultrasound will substantially increase the level of nodular thyroid disease detected in a population compared to that found by physical examination (palpation) alone. Similarly, different thresholds for laboratory values used to define a case of hypothyroidism could result in apparent differences in disease occurrence that simply reflect differences in diagnostic definition. Such detection effects can be substantial. For example, there is direct evidence in the HTDS of a large "screening effect" for thyroid cancer. Twelve of the 20 cases of thyroid cancer among the 3440 evaluable study participants were detected as a result of the HTDS examinations, and 2 of the 12 cases were diagnosed by palpation only after the ultrasound scan was reviewed and the participants were re-examined. The resulting cumulative incidence for thyroid cancer was 2.5 times greater than what it would have been had it been based on cases identified through the normal medical care system.

Third, populations with different characteristics or different exposures which might affect the occurrence of thyroid disease can exhibit very different disease rates. For example, rates of most forms of thyroid disease are higher for females than males, and increase with increasing age. Thus, all other factors being comparable, two populations with different age and gender structures might exhibit very different rates of thyroid disease. Similarly, people living in an iodine deficient environment would likely have different rates of some forms of thyroid disease than people who are iodine sufficient.

Despite these substantial obstacles to making valid comparisons between the cumulative incidence of specific outcomes determined in the HTDS and estimates found in the published literature, we attempted

to assemble the most comparable information possible for the most important outcome categories studied. A summary of this information is presented in the subsections below, with special attention given to differences between specific studies and the HTDS which could account in part or entirely for differences in reported disease rates. Although admittedly imperfect, these data provide at least a frame of reference within which the cumulative incidence data from the HTDS can be evaluated.

E.1. Prevalence of Thyroid Cancer

The occurrence of thyroid cancer varies widely worldwide, is more common among females, and increases sharply with increasing age. Annual incidence rates have been reported to range from a high of 104 cases per million in women in Hawaii to a low of 14 cases per million in women in Poland (140). The age-adjusted annual incidence in the United States is 55 cases per million people (80 per million in women and 29 per million in men)(140). Further, the incidence of thyroid cancer in the United States has steadily increased over the last several decades, perhaps in part due to improved methods of diagnosis (140). Although it might be preferable to compare the occurrence of thyroid cancer in the HTDS cohort to that in other populations using incidence data, the retrospective nature of the HTDS design precluded us from accurately determining a date of diagnosis for each case, and therefore from calculating an incidence rate in the cohort.

It is possible to use incidence data to predict the cumulative incidence of thyroid cancer that might be expected in the HTDS cohort, although such predictions must be interpreted cautiously. In fact this was done at the beginning of the study to assist in developing the study design. As described in Appendix H of the HTDS Protocol (1), the cumulative incidence of thyroid cancer for the HTDS cohort was estimated using age- and sex-specific incidence rates from the Cancer Surveillance System (CSS), a population-based registry for the thirteen northwestern counties of Washington State. To account for the screening effect of the HTDS clinical examinations, the CSS incidence rates were multiplied by three, using a value suggested by the National Council on Radiation Protection and Measurements (141). The predicted cumulative incidence of thyroid cancer was 0.0068 (0.68%) for women and 0.0025 (0.25%) for men. These predictions are in good agreement with the observed values of 0.7% for women and 0.4% for men (see Section IX.C above). While this may be viewed as evidence that overall thyroid cancer rates for the HTDS cohort are not higher than expected, it must be recognized that incidence rates in the population covered by the CSS may differ from the background rates of the HTDS cohort, and that the factor of three assumed for the screening effect may not be appropriate for the HTDS clinical evaluation.

Unfortunately, there is very little information available regarding the prevalence of thyroid cancer in the general population. This is due primarily to two reasons. First, because the absolute frequency of incident cases is quite low, screening programs in the general population are not very feasible and are generally not considered an appropriate use of resources. Second, it can be difficult to discriminate between clinically significant thyroid cancer and that which does not adversely impact a person's health. The latter is usually referred to as occult or microscopic cancer.

There have been a number of studies of patients with thyroid nodules who are referred for surgery (142). Very high prevalence rates of thyroid cancer (5-24%) have been reported from these studies. However, such surgical series have a high likelihood of selection bias since such patients are usually referred because of high suspicion for thyroid cancer. Consequently, these studies almost certainly overestimate the true prevalence of thyroid cancer in the general population.

The best data for estimating the frequency of "occult" or microscopic thyroid cancer come from autopsy studies, where microscopic thyroid cancer is found in people who died of other causes. Crapo and Wang summarized a series of nine autopsy studies, performed from 1952-1977, which showed a mean prevalence of thyroid cancer to be 3.6% among 3744 cases (range of prevalence 0.45-13.0%) (142). These studies were chosen in part because they all were carefully performed, each examining 1-3 mm slices of thyroid tissue.

In contrast to clinically important thyroid cancer, most studies show that occult thyroid cancer does not seem to vary by age or gender. These studies also show that the correlation between prevalence of

occult thyroid cancer and mortality is poor. Countries such as Japan have a high prevalence of occult thyroid cancer but low mortality, whereas other countries have both low prevalence and low mortality. For example, Fukunaga reported a 24-28% prevalence of occult thyroid cancer from autopsy studies of Japanese and yet there is a low mortality rate from thyroid cancer in Japan (143).

In summary, there are no good estimates of thyroid cancer prevalence to which the cumulative incidence findings in the HTDS cohort can be compared. The prevalence estimates that are available are most certainly overestimates of what might reasonably be expected in the HTDS cohort, as they are derived from either patients referred for surgery, or from autopsy studies of occult cancer.

E.2. Prevalence of Thyroid Nodules

It is well known that thyroid nodules are a common finding in the general population (reviewed in 142). The primary determinant of the variation in prevalence estimates of thyroid nodules is the method of detection. Estimates vary widely, depending on whether the method of detection is palpation, ultrasound, or autopsy.

The oldest and most widely quoted study of thyroid nodularity in the general adult population is the Framingham Study, which began in 1948 and employed palpation as the method of detection (144). The initial cohort was composed of 5127 randomly selected individuals from the town of Framingham, Massachusetts who were given careful thyroid physical examinations to determine the prevalence of thyroid nodules. The age range was 30-59 and the geographical area was not felt to be iodine deficient. The criteria for a definite solitary thyroid nodule was one that was palpable by at least two examiners, while suspected nodules were those palpable by only one examiner. The average diameter of nodules was 1 cm. The prevalence of definite single nodules detected over a 5 year examination period was 1.9% (2.7% for females; 0.8% for males), while the combined prevalence of definite and suspected solitary nodules was 3.0% (4.6% for females, 1.1% for males). An additional 1.1% of the cohort had multiple palpable nodules (1.7% for females, 0.4% for males). Thus, of the total 5127 people examined, 218 people, or 4.2% had palpable thyroid nodules (6.4% for females, 1.5% for males).

A 15-year follow-up study of this cohort was subsequently published in 1968 (145). Of the 218 people found to have thyroid nodules in the initial survey, 139 people still had nodules which were unchanged at the 15-year follow-up. Of the remaining 79 people, 45 had nodules excised during the follow-up period (all were benign), 15 had died (none of thyroid related causes), and 19 had nodules that were excised prior to the initial survey (all benign). Of 4909 people who were free of palpable thyroid disease at the initial survey, 67 people (1.4%) developed new nodules during the 15 year follow-up. Although none of these new nodules were reported to have thyroid cancer, only 13 people actually had surgery; the remainder were thought to be clinically benign.

Thus, the cumulative incidence of palpable thyroid nodules at the end of the 15 year follow-up period in the Framingham Study was 5.6% (285/5127 people); for females the cumulative incidence was 8.1% (230/2845) and for males 2.4% (55/2282). Of the total 285 people with nodules, all were thought to be clinically benign. Although only 27% had surgical excision, none showed any evidence of malignancy**.** The initial study attempted to discriminate between solitary nodules and multiple nodules (73% were solitary), however the follow-up study did not and included all nodules in the prevalence data whether they were thought solitary or multiple. These estimates of nodule prevalence are probably the most comparable to the HTDS experience found in the world literature: they represent reasonably long-term follow-up, the age range at the end of follow-up is approximately 45-74, most people were examined by multiple physicians, the estimates include people with prior surgery, and the population under study is a randomly selected group.

A similar study by Whickham et al. documented the prevalence of thyroid disorders in 2779 adults who were age and sex matched to the British population (146, 147). Although this study provides some of the highest quality information regarding thyroid dysfunction and autoimmune thyroid disease in an unselected population (see below), the results of the 20 year follow-up study published in 1995 regarding

the prevalence of thyroid nodules is of limited value because the study was not designed to assess nodular thyroid disease. The original Whickham study published in 1977 contained one brief statement that thyroid nodules were detected in 5.3% of women and 0.8% of men. However, no details regarding the characterization of nodule size or data regarding thyroid cancer in this cohort were reported (146).

There are no other population-based studies of thyroid nodularity in adults. Ezzat (99) reported results from a small series of adult volunteers who responded to an employee bulletin board and who were given thyroid examinations by two examiners as well as ultrasound and laboratory evaluations. Of the 100 people participating, 21 (21%) had palpable nodules which were confirmed by ultrasound. The high prevalence is likely influenced by the predominance of females (84%), and perhaps the wide age range of 25-77 years.

The use of ultrasonography has greatly increased the sensitivity of detecting anatomical abnormalities or variations of normal in the human thyroid. This technology has raised an important question of whether the high frequency of thyroid UDAs in the general population constitute clinical disease or whether many of these abnormalities represent variations in anatomy that do not adversely affect health. Section X.E.6 below considers the prevalence of thyroid UDAs.

Estimates of the prevalence of thyroid nodules have also been reported based on autopsy findings. One of the most quoted studies of thyroid nodules detected in people dying of non-thyroid disease was published in 1955 by Mortenson (107). These authors performed 821 consecutive autopsies (age range 0-99) and found a prevalence of thyroid nodules to be 49.5%. An older study (147) in 1938 (age range 0-89) found a prevalence of only 8.2% but limited their findings to nodules greater than 2 cm. In 1965 Oertel reported a thyroid nodule prevalence of 13% in previously healthy military men dying from non-thyroid causes (n=113) (149). Other studies by Rice and Hull have found an even higher prevalence of nodules at autopsy (57% and 65% respectively), but were conducted in endemic goiter areas which might explain the high rates of nodularity (150, 151).

In summary, the prevalence of thyroid nodules identified in the HTDS cohort (7.2% overall; 9.7% in females and 4.7% in males) is similar in magnitude to that found in the two population-based studies reported in the literature. The slightly lower prevalence in the Framingham cohort most likely reflects a considerably shorter period of follow-up, younger age range, and the absence of the ultrasound screening effect demonstrated for benign nodules in the HTDS. The latter effect of excluding ultrasound has been demonstrated in the HTDS cohort: the cumulative incidence of benign nodules by palpation only (not influenced by ultrasound) is 4.8% for females and 2.0% for males. This value is in quite good agreement with the Framingham prevalence figures. Estimates of prevalence of thyroid nodules based solely on ultrasound detection, or autopsy findings, are considerably higher than in the HTDS cohort and should probably be regarded as an indication of the upper bound of possible prevalence in human populations.

E.3. Prevalence of Hypothyroidism

Hypothyroidism is generally classified into two categories based on severity. Overt hypothyroidism usually produces symptoms and is diagnosed by both elevated TSH levels and decreased levels of circulating thyroid hormone. Subclinical hypothyroidism may or may not produce overt symptoms. It is generally agreed that subclinical hypothyroidism is present when the TSH is between 5 and 10 μIU/ml and thyroid hormone levels are normal. The degree to which subclinical hypothyroidism is included in prevalence studies of hypothyroidism can greatly influence the magnitude of the estimates. In addition, age, gender, and the presence of iodine deficiency or autoimmune thyroid disease also influence the magnitude of the prevalence estimates.

Perhaps the most thorough evaluation of the prevalence of hypothyroidism in an unselected population is the Whickham study and its 20-year follow-up study (146, 147). In a review by Wang and Crapo, they indicate it is “the only study that has surveyed a representative sample of the entire adult population of a large community for thyroid disease by employing detailed medical histories, rigorous physical examination, and sophisticated laboratory testing” (142).

The Whickham survey sample was randomly selected from an electoral register of adults older than 18 years in Great Britain. Of the initial sample of 3538 people, 2779 people participated in the study, 1285 men and 1494 women. The age, sex, and social class of the sample generally reflected that of Great Britain. In addition to detailed history and physical examination, each participant was tested for TSH by RIA, free thyroxine index, antithyroid thyroglobulin antibody, and AMA (antithyroid microsomal antibodies). The prevalence of subclinical hypothyroidism (TSH>6 μIU/ml) was 7.5% in women and 2.8% in men (combined, 5.3%). The TSH levels increased with age in women but not in men. The increase with age in women was also seen primarily in those women with positive antithyroid antibodies.

One of the important aspects of the Whickham cohort is that it was followed for 20 years to evaluate the natural history of thyroid disease. Some type of follow-up information was available in over 95% of the original cohort. The results showed that, after a median follow-up of 19 years, the prevalence of hypothyroidism increased significantly. This was in contrast to hyperthyroidism, which remained almost unchanged. The prevalence of spontaneous hypothyroidism in the cohort at the end of follow-up was 4.7% (7.7% in women and 1.3% in men). The overall median age was 58 (38-93), with a median age of 58 for men and 59 for women. These numbers increased further with older women with a prevalence of 10.4% for women older than 45 and 17.5% for women older than 75.

Sawin and colleagues evaluated the Framingham cohort in 1985 and assessed the frequency of hypothyroidism (152). The age range of this cohort, begun in 1948, was between 60 and 89 years of age. For the total cohort of 2139 people at the end of the 15-year follow-up period, 10.3% had an elevated TSH (13.6% for women, 5.7% of men). Excluding those with subclinical hypothyroidism (TSH 5-10 μIU/ml), those with clearly elevated TSH levels (>10 μIU/ml) included 4.4% of the total cohort (5.8% for women and 2.4% in men). Thus, in an unselected aging population, the total prevalence of hypothyroidism was quite high at over 10%.

More recently (1993), Geul and colleagues conducted a population survey in the Netherlands which corroborates the Whickham results regarding risk factors for progression of hypothyroidism (10188). TSH and AMA were measured in 423 randomly selected women from the Netherlands, age 40-60, and were repeated ten years later. The prevalence of hypothyroidism at the end of 10 years in the cohort was 7.3% for women, mean age 65.

Several other studies have investigated thyroid deficiency in population based settings. One recent study (1999) evaluated 1411 people representing the majority of individuals from the population of Pescopagano, an iodine deficient community in southern Italy (153). This cohort represented a relatively young population with only 28% of people older than 46 and 30% younger than age 15. Overt hypothyroidism occurred in only 0.2% whereas subclinical hypothyroidism (TSH > 3.7 μIU/ml) occurred in 3.8%. Although this was reported not to be significantly different from that reported in the Whickham study, it is lower than a number of other reports including Framingham and likely reflects the relatively young age of the cohort and perhaps iodine deficiency in the population.

A population-based survey of Danish centenarians has provided interesting information about the effect of very old age on thyroid dysfunction (154). A total of 140 people older than 100 years agreed to have blood tests taken. The number of people with subclinical hypothyroidism was fairly small at 2.9%. An additional 2.9% reported previous hypothyroid disease. The authors concluded that the level of thyroid dysfunction in people older than 100 years was not significantly increased over older people younger than age 100, and that thyroid function in centenarians was well preserved.

The recently published Colorado Health Study involved screening of thyroid function in over 25,000 people at a Health Fair (155). The mean age of the group screened was 56 years with women representing 56%. An elevated TSH was detected in 2450 people (9.5%). Of this group, 1799 people (7.0%) had subclinical hypothyroidism (TSH between 5.1-10 μIU/ml) and 619 people (2.4%) had a TSH greater than 10 μIU/ml. For the age group 45-54, the prevalence of elevated TSH levels was 5% for males and 9% for females; for the 55-64 age group the prevalence increased to 6% for males and 13% for females and continuing increasing with age to about 21% for women greater than age 74.

Finally, Hollowell and coworkers recently reported the results of a large screening study of thyroid abnormalities in a population sampled to represent the geographic and ethnic distribution of the US population (156). The cohort consisted of 31,000 people age 6 and older. The mean TSH was 1.49 μIU/ml for those above 12 years who did not report thyroid disease or thyroid medication. For the age range 40-49, the percentage of people with TSH above 4.5 μIU/ml was 5.7% for females and 3.7% for males, whereas the frequency of positive TPO antibodies was 17.2% in females and 11.3% in males. For the age range 50-59, TSH was greater than 4.5 μIU/ml in 8.1% of females and 2.4% of males; the prevalence of positive TPO antibodies was 18.2% in females and 10.5% in males. While the prevalence of elevated TSH levels was greater among females than males, this difference was not significant after controlling for TPO antibodies. This result is consistent with the results of the Whickham and Geul studies reported above.

In summary, there is considerable information available in the published literature regarding the prevalence of hypothyroidism (both overt and subclinical) in a number of population-based samples of individuals. Estimates of prevalence from the major studies are in reasonable agreement with each other, and define a range which encompasses the estimates derived in the HTDS (7.8% overall; 11.7% in females and 3.7% in males).

E.4. Prevalence of Autoimmune Thyroiditis

Estimating the prevalence of autoimmune thyroiditis is particularly challenging because the antibody assays for detecting autoimmune thyroiditis have changed over time. These assays have ranged from antithyroglobulin measurement via agglutination techniques to antithyroid antimicrosomal antibodies to current and refined methods for detecting thyroid peroxidase antibodies. The reported prevalence of autoimmune thyroiditis in any given study depends on a number of factors, but especially the type of assay used.

The prevalence of antibody positivity in the general population is generally much higher than the prevalence of clinical disease. Although the ability to detect individuals with positive antithyroid antibodies has greatly enhanced the ability to predict risk for developing hypothyroidism, it is nevertheless difficult to predict which individuals with antibody positivity will develop clinical disease. In part, the probability of developing disease is related to the magnitude of the positive test. Summarized below are results from studies of the prevalence of antibody positivity in the general population, with an emphasis on those studies in the last 10-15 years which have utilized more highly sensitive antibody assays.

The Whickham study (reviewed above) also provides important results regarding autoimmune thyroiditis. At the 20-year follow-up, 19% of the cohort had positive antithyroid antibodies (147). The prevalence in women was 26.4% and in men 8.8%. These antibodies were later reported by the authors to be TPO antibodies (10332).

An important study by Spencer and colleagues evaluated antibody positivity in thyroid cancer patients and compared them to a group of 4453 people representing the general population who were undergoing routine multiphasic health examinations (157). The mean age of the healthy participants was 45 with a range of 12-99, and a male to female ratio of 0.69. Antibodies to both thyroid peroxidase (TPO) and thyroglobulin were measured. The prevalence of anti-TPO alone was 4.0%, anti-TG was 3.1%, and both TPO and TG antibodies was 7.0%. The prevalence of having <u>any</u> antibody positivity was 14.1%.

The Pescopagano study (described in section X.E.3 above) also assessed antithyroid antibody positivity. The overall prevalence of people positive for both TPO and TG antibody tests was 12.6% (females 17.3%; males 7.0%). Positive antibody tests showed an age effect with a prevalence of 2.4% in children, increasing to 22% in people aged 46-55. No further increases were seen in older people. Although low titer antibody positivity was quite frequent in this cohort, the authors concluded that the spectrum of thyroid disease was not different from that observed in iodine-sufficient areas.

Five additional studies provide data on the prevalence of anti-thyroid antibodies in a healthy unselected population. The Geul study from the Netherlands (discussed in section X.E.3 above) showed that the progression to hypothyroidism was strongly influenced by the presence of autoimmune thyroiditis (158). In a group of 427 women with mean age of 55 (40-60), AMA (antithyroid microsomal antibodies) were measured at the start of the study. The prevalence of positive AMA was 11%. Prentice measured TPO antibodies in 698 female blood donors from seven towns in Great Britain and reported that 18% were positive. The prevalence rose from 15% in women of age 18-24 to 24% in women of age 55-64 (159). Lazarus found similar results in screening 414 asymptomatic elderly women over age 70. The prevalence of positive TPO antibodies was 15% and anti-thyroglobulin antibodies 13% (160). An even higher frequency of positive TG antibodies was found in a small study of patients with thyroid disease compared to 140 healthy volunteers. The volunteer group consisted of 80 women (median age 50) and 60 men (median age 48) in whom care was taken to exclude the presence of thyroid disease. In this group, 27% had positive TG antibodies (161). This is one of the highest frequencies of positive thyroid antibodies reported in a healthy population, although these results are somewhat limited by a very small number of people screened. Finally, the Danish centenarians (described in section X.E.3 above) were also assessed for antithyroid antibodies (154). They were classified into two groups; dependent or independent based on their need for assistance for daily living activities. Those classified as dependent had higher rates of positivity (TPO 11.1%; TG 14.8%; both 22.2%) than those classified as independent (TPO 6.8%; TG 5.1%; both 8.5%).

In summary, there is also considerable information available regarding the prevalence of positive anti-thyroid antibodies in the general population. Estimates range from 3-27%, but are highly variable and are dependent on a number of factors including age, gender, geographical location, type of antibody assay, and perhaps ethnic background and iodine sufficiency as well. Nevertheless, the cumulative incidence estimates for autoimmune thyroiditis in the HTDS cohort (18.2% overall; 23.1% in females and 13.1% in males) are consistent with these estimates in other populations.

E.5. Prevalence of Hyperparathyroidism

The parathyroid glands, located in the back of the thyroid gland, contribute to the regulation of calcium levels in the body through the production of parathyroid hormone (PTH). The most common parathyroid disorder is hyperparathyroidism, which results in high circulating calcium levels due to high levels of PTH secreted from one or more of the parathyroid glands. This disorder is uncommon in comparison to thyroid disease.

Primary hyperparathyroidism has traditionally been defined as an elevated calcium level in the presence of an elevated PTH level. However, because accurate tests of PTH have only become available in the last 10-15 years, early studies have used variable definitions for the disease. Even with the accurate PTH tests available today, the frequency of hyperparathyroidism found from one study to another depends greatly on the cut-off points used by the investigators. For example, a high normal PTH level in the presence of a high calcium is not truly normal and usually represents primary hyperparathyroidism. However, differences in the actual cut-off used from one study to another will result in variable prevalence rates being reported across populations.

Early studies reported a prevalence of hyperparathyroidism of between 0.29% and 1.03% in Swedish men and women age 50-63 (162), and a prevalence of 1.5% in women older than 60 (163). However, these estimates did not reflect a sample of the general population. The prevalence of hyperparathyroidism in the unexposed control group from the studies of atomic bomb survivors in Japan has been reported to be 0.1% in men and 0.3% in women over age 41 (164).

Lundgren and colleagues evaluated 5202 women attending a population-based mammography screening program in Sweden (165). Several definitions of hyperparathyroidism were employed which varied the cutoff points of calcium and PTH. The prevalence of primary hyperparathyroidism in women age 55-75 was 2.1%. This prevalence exceeded that reported by Christensson, which required indisputable hypercalcemia (greater than 2.78 mmole/L) for the diagnosis of hyperparathyroidism rather than high

normal calcium levels in the presence of an increased PTH. Lundgren concluded that the use of current biochemical criteria results in under-diagnosis of primary hyperparathyroidism.

A recent population-based study by Jorde and colleagues from Norway measured serum calcium in approximately 25,000 people who participated in a broad health survey (166). In people with calcium levels greater than 2.59 mmol/L, PTH was also measured. The prevalence of primary hyperparathyroidism in this group (ages 25-75) was 0.17% for men and 0.45% for women ($p<.001$). A subgroup analysis was performed in older women between ages 50 and 75. Using the criteria for the main study, the prevalence of hyperparathyroidism was 8.8%. However, the prevalence varied dramatically from 3.6% to 13.9% when the criteria for hyperparathyroidism were varied.

Despite the substantial difficulties in comparing prevalence estimates of hyperparathyroidism in different populations due to differences in diagnostic definitions used, the cumulative incidence estimates from the HTDS (0.3% overall; 0.6% in females and 0.1% in males) are well within the range of estimates found in the published literature.

E.6. Prevalence of Thyroid Ultrasound-Detected Abnormalities of the Thyroid (Thyroid UDAs)

During the last 15 years, high-frequency ultrasound has increasingly been used in the evaluation of thyroid nodules. Although the traditional definition of a thyroid "nodule" has been based on clinical palpation, the greater sensitivity of ultrasonography has led to its increased use and consequently the detection of nonpalpable, millimeter-size abnormalities. This has raised several important issues: 1) thyroid UDAs have been shown to occur frequently in the general population, without an adequate understanding of their risk of malignancy or biologic significance; 2) thyroid UDAs have often been classified as "nodules" regardless of size. This has resulted in uncertainty about whether palpable ultrasound-detected nodules are biologically different than the large numbers of nonpalpable ultrasound-detected "nodules"; 3) the use of ultrasound in defining criteria for thyroid nodules has made it difficult to compare clinical thyroid outcomes across epidemiological studies if they use different criteria for thyroid nodularity; and 4) although ultrasound has exceptional sensitivity, recent data regarding specificity (the ability to distinguish benign from malignant nodules) suggests that the increases in specificity of ultrasonography are associated with significant decreases in the sensitivity.

A number of studies have shown that the prevalence of thyroid UDAs is high in the general population. Tan et al. have recently reviewed the literature and reported a range of prevalence of 17-67% (97). In 1000 people evaluated for hypercalcemia, in whom 8% had a nodular goiter, 46% had discrete lesions on ultrasound and 38% were reported to have nodules (98). The study reporting the highest prevalence of thyroid UDAs was a prospective study of 100 employees responding to a notice on a bulletin board: 67% of these women, mean age 43, showed abnormal thyroid ultrasound scans (10114). Thyroid UDAs in populations without apparent thyroid disease have also been documented outside the US with prevalence figures ranging 17-27% (100, 101, 102). Most of these studies have been consistent in showing that nonpalpable thyroid UDAs are generally small and that solitary nodules on clinical examination are often associated with multiple other thyroid UDAs. Both Tan (103) and Brander (100) have demonstrated that in patients with known palpable thyroid nodules greater than 1 cm, 48% harbored additional thyroid nodules found on ultrasound.

Brander and colleagues have also published a comprehensive study regarding the prevalence of thyroid UDAs. They randomly selected 253 people from a Finnish city council registry and screened for thyroid UDAs (104). The sample was distributed evenly among four age brackets from 20 through 50. The community was not thought to be endemic for goiter. Thyroid UDAs were detected in 69 people (27.3%). These abnormalities were solitary in 57%, multiple in 22%, and diffuse in 22%. The mean age for people with normal ultrasound scans was 35, the mean age for the group with abnormal ultrasound findings was 37. These abnormalities were found more often in women than men and increased with age for both sexes. For women, the prevalence of thyroid UDAs was 30% in the 20-29 age group, 32% in the 30-39 age

group, and 41% in the 40-50 age group. All participants underwent thyroid palpation prior to ultrasound examination. Palpable abnormalities were detected in 13 people (5.1%); three with a solitary nodule, five with multiple nodules, and five with abnormal consistency. Fine needle aspirations were done in 30 individuals. All were negative for malignancy with one intermediate probably of neoplasm; that person underwent surgery and had a follicular adenoma.

Bruneton evaluated 1000 healthy volunteers without history of thyroid disease and performed high frequency thyroid ultrasound examinations (106). Although selection criteria or mean age were not provided, 57% of participants were over 50 years. Ultrasonography was performed with 13 MHz transducers and all ultrasound nodules greater or equal to 3 mm were counted. One or more nodules were detected in 34.7% of subjects. For people less than age 50 (n=431), the prevalence was 25%. For people greater than age 50 (n=569), prevalence was 42%. For all ages, the prevalence in women was 44% and the prevalence in men was 17.7%.

A Belgian study assessed thyroid ultrasound abnormalities in 300 patients who were referred for abdominal ultrasound examinations (102). Although this study sample is not a random representation of the general population, there were extensive exclusion criteria for those with symptoms or signs of thyroid disease. Unlike the Bruneton study, this investigation used a 5.5 MHz ultrasound transducer. The mean age was 47 (range 1-88 years) and 55% of the participants were males. Small echoic nodules were found in 19% of patients. In patients in their seventh decade of life, the prevalence increased to over 40%. The wide age distribution of this cohort and the high percentage of males undoubtedly influenced these results.

In summary, there is considerable published evidence reporting high prevalence of thyroid abnormalities detected by ultrasound examination in the general population. Estimates of 40%-50% or even greater are not uncommon, depending upon the characteristics of the population screened and the technology used. The prevalence of any UDA found in the HTDS (46.5% overall; 55.5% in females and 37.4% in males) are consistent with these estimates.

In overall summary, a considerable effort was made to assess the world literature on the prevalence of the major thyroid and parathyroid disease outcomes evaluated in the HTDS as well as thyroid UDSs. This was done in order to compare the disease experience of the HTDS cohort to what might reasonably be expected based on the experience in other populations not exposed to Hanford radiation. As outlined at the beginning of this section, comparisons of this type are imperfect and must be interpreted with great caution. What appear to be differences in prevalence estimates between the HTDS cohort and other populations may well reflect differences in any of a number of factors other than exposure to radiation from Hanford. Nevertheless, it is clear from comparisons with the most comparable studies in other locations that for the major outcomes described above (thyroid nodules, thyroid cancer, hypothyroidism, autoimmune thyroiditis, hyperparathyroidism, and thyroid UDAs), the estimates of cumulative incidence or prevalence derived from the HTDS are well within the range and are consistent with published estimates. There is no indication that the levels of disease occurrence in the HTDS cohort are systematically different, or higher, than what has been reported around the world in a variety of different circumstances.

F. Summary and Conclusions

The HTDS was conducted to determine whether exposure to atmospheric releases of primarily ^{131}I from the Hanford Nuclear Site between 1944 and 1957 resulted in increased thyroid disease among those exposed. The study evaluated twelve categories of thyroid disease, the results of several laboratory tests for thyroid function, anti-thyroid antibody and serum calcium level, thyroid UDAs, thyroid mass, and hyperparathyroidism. The primary analysis utilized an estimate of thyroid radiation dose for each individual based on information about their residence history and dietary consumption patterns during the times of the Hanford releases. Additional analyses were also conducted using several alternative methods for estimating dose, both quantitative and qualitative, including methods which were independent of the HEDR models. The primary analyses were based on a sex-stratified linear dose-response model, although alternative models for the shape of the dose-response were also investigated. The potential effect on any

dose-response of a number of lifestyle factors and indicators of other radiation exposure were evaluated as covariates in the models. All primary dose-response analyses were repeated to include adjustments for uncertainty in the individual radiation dose estimates.

This study found no statistically significant association between dose to the thyroid from Hanford radiation and 1) cumulative incidence of any of the disease outcomes; 2) prevalence of thyroid UDAs; or 3) thyroid laboratory tests or thyroid mass. There was also no evidence of a dose-response for hyperparathyroidism, although a positive dose-response was seen for serum calcium. An increasing thyroid dose was significantly associated with a decrease in serum calcium. Although the explanation for this result is not clearly apparent, the finding does not appear to be of clinical significance (discussed more fully in section IX.Q.7 above). These results remained the same when alternative methods of assessing radiation dose were used, and after accounting for uncertainty in dose estimation. Based on data available regarding the tracing and enrollment of study participants, there is no evidence that the absence of a dose-response relationship is due to bias in selection of the cohort, loss to follow-up, or enrollment and participation.

Given the principal differences between the radiation exposure circumstances at Hanford and those of other populations studied in relation to radiation-induced thyroid disease (summarized above), the findings of this study are not inconsistent with the current published literature regarding the effect of exposure to radioactive iodine and the risk of thyroid and parathyroid disease. This is particularly so given the relatively small magnitude of the estimated thyroid radiation doses in members of the HTDS cohort (mean = 174 mGy) and the relatively protracted nature of the exposure over time. There is little evidence in the literature to suggest that people exposed to ^{131}I at the levels found in this study over a period of months or years would experience higher rates of thyroid or parathyroid disease as a result of their exposure.

Nevertheless, a lingering question for many may be whether the uncertain nature of the dose estimation used in the primary analyses is so great that it renders the quantitative dose-response results inconclusive. The study has attempted to address this possibility in three ways. First, alternative qualitative methods of assigning exposure were used. Results from these analyses were consistent with those from the quantitative dose-response analyses. Second, two different approaches were employed to evaluate the impact of dose uncertainty on the primary risk estimates. Neither resulted in findings that were materially different from those ignoring such uncertainty. Third, the impact of dose uncertainty on study power was assessed using simulation methods. These analyses revealed that the reduction in statistical power due to uncertainty in dose estimation was modest, and that even after accounting for such uncertainty the study had adequate statistical power to detect effects as small or smaller than those in the existing published literature. Although any epidemiologic study is limited to some extent by uncertainty in the assessment of exposure, the impact of such uncertainty on the power of the study and the estimation of risk is seldom addressed to the extent attempted here. Further, the fact that epidemiologic investigations are inherently "uncertain" does not imply complete randomness or unpredictability, nor does it mean that reasonable conclusions cannot be drawn from such studies.

In conclusion, the results of the HTDS provide no evidence of a statistically significant association between increasing thyroid radiation dose from Hanford and the cumulative incidence of any of the disease outcomes studied or the prevalence of thyroid UDAs. These findings do not definitively rule out the possibility that Hanford radiation exposures are associated with an increase in one or more of the outcomes under investigation. However, it does mean that if such associations exist, they were likely too small to detect using the best epidemiologic methods available.

REFERENCE LIST

1 Hanford Thyroid Disease Study Protocol, May 23,1993

2 Hanford Thyroid Disease Study Pilot Study Final Report. Fred Hutchinson Cancer Research Center, January 25, 1995

3 U.S. Department of Commerce, U.S. Bureau of the Census. Census of Population: 1960. Vol. 1. Characteristics of the Population. Part 49. Washington

4 Cantrell, S. and Healy, J., "Iodine Metabolism with Reference to I-131"; HW-7-2604; (Richland, WA: Hanford Engineer Works, 10/22/45) and Thornburn, C. "Detection of Plutonium in Flora"; HW-7002; (Richland, WA: Hanford Engineer Works); and Richland Villager, 2/27/47, p.1, and 3/20/47, p.1, and 4/3/47, p.1, and 8/7/47, p.1.

5 "Tolerance Limits;",HW-72602 (Richland, WA: Hanford Engineer Works, 1945 Oct 18); and Cantrell, S. and Healy, J., ; HW-7-2226-4; and Parker, H. M. "Proposed Revisions of Tolerance for I-131," HW-7-3042 (Richland, WA: Hanford Engineer Works, 1945 Dec 17).

6 Herde, K., I-131 Deposition in Cattle Grazing on North Margin of HEW," HW-33628 (Richland, WA: Hanford Engineer Works, 1946 Aug 29); and Herde, K., "I-131 Accumulation in the Thyroid of Sheep Grazing Near HEW," HW- 33455 (Richland, WA: Hanford Engineer Works, 1946 Mar 1); Libby, L. M; The Uranium People; (New York; Crane, Russak, Inc.; 1979) p.174.

7 Parker, H.M., "Subject: Radiation Exposure Data," Letter To: Sturges, D.G. (Chief, Operations Division, AEC-Hanford Operations Office). HW-19404, Richland, WA: Hanford Engineer Works; 1950 Nov 80, p.11.

8 Roberts, R. E., "History of Airborne Contamination and Control - 200 Areas," HW-55568-HD (Richland, WA: Hanford Atomic Productions Operation; 1958 Apr 1), pp. 2-3,6.

9 Paas, H. J. and Singlevich, W. "Radioactive Contamination in the Environs of the Hanford Works, over the Period October, November, December 1949," HW-17003 (Richland, WA: Hanford Engineer Works; 1950 Feb 3), pp.12-15,18,21-22,26,31.

10 Harlan, W. F.; Jenne, D. E., and Healy, J. W. Dissolving of 20-Day Metal at Hanford. HW-17381 (Richland, WA: Hanford Engineer Works; 1950 May 1).

11 Mooney, R. R. Special Report: Preliminary Dose Assessment of Hanford Historical Releases 1944-1956. Department of Social and Health Services, Office of Radiation Protection, Olympia, WA, 1986.

12 Cate, S., Ruttenber, A.J., Conklin, A.W. The feasibility of an epidemiologic study of thyroid neoplasia in persons exposed to radionuclides from the Hanford nuclear facility between 1944 and 1956. Health Physics. 1990; 59: 169-178.

13 Conklin, A. W. Releases of Radioactivity from Hanford, 1944-1956. Memorandum dated 1987 Jul 1. Department of Social and Health Services, Office of Radiation Protection; Olympia, WA.

14 Hanford Environmental Dose Reconstruction Project. Draft Air Pathway Report. PNL-7412 HEDR. Battelle Pacific Northwest Laboratories. Richland, WA, 1990.

15 Smeesters, P, J Fruhling, L Van Bladel, A Wambersie. [Nuclear accidents and iodine prophylaxis. Part 1: Risks due to irradiation of the thyroid gland]. [French]. Revue Medicale de Bruxelles 19[6], 475-82. 1998.

16 Shipler, D, B Napier, W Farris, MD Freshley. Hanford Environmental Dose Reconstruction Project - An overview. Health Physics 71[4], 532-544. 1996.

17 Farris, WT// Napier, BA// et al. Radiation doses from Hanford Site releases to the atmosphere and the Columbia River. Health Physics; 1996; 71 (4): 588-601

18 Farris, WT// Napier, BA// Eslinger, PW//et Atmospheric Pathway Dosimetry Report, 1944-1992 PNWD-2228 HEDR

19 Nagataki, S.; Shibata, Y.; Inoue, S.; Yokoyama, N.; Izumi, M., and Shimaoka, K. Thyroid diseases among atomic bomb survivors in Nagasaki [published erratum appears in JAMA 1995 Jan 25;273(4):288]. JAMA. 1994 Aug 3; 272(5):364-70; ISSN: 0098-7484.

20 Maxon, HR; Thomas, SR; Saenger, EL, and et al. Ionizing irradiation and the induction of clinically significant disease in the human thyroid gland. American Journal of Medicine. 1977; 63(6):967-78.

21 Duffy, F. Cancer of the thyroid in children: A report of 28 cases. J Clin Endocrinol Metab. 1950; 10:1296-1308.

22 Clark, D. E. Association of irradiation with cancer of the thyroid in children and adolescents. JAMA. 1955; 159:1007-1009.

23 Simpson, C. L.; Hemplemann, L. H., and Fuler, L. M. Neoplasia in children treated with x-rays in infancy for thymic enlargement. Radiology. 1955; 64:840-845.

24 Beach, S. A. and Dolphin, G. W. A study of the relationship between x-ray dose delivered to the thyroids of children and the subsequent development of malignant tumors. Phys Med Biol. 1962; 6:538.

25 Kaplan, M. M.; Garnick, M. B.; Gelber, R.; Li, F. P.; Cassady, J. R.; Sallan, S. E.; Fine, W. E., and Sack, M. J. Risk factors for thyroid abnormalities after neck irradiation for childhood cancer. American Journal of Medicine. 1983 Feb; 74(2):272-80; ISSN: 0002-9343.

26 Ron, E. and Modan, B. Thyroid cancer and other neoplasms following childhood scalp irradiation. IN Radiation Carcinogenesis: Epidemiology and Biologic Significance. Boice, J. D. and Fraumeni, J. F. Jr., Editors. New York, New York: Raven Press; 1984; p. 139.

27 Schneider, A. B.; Shore-Freeman, E.; Ryo, U. Y, et al. Radiation-induced tumors of the head and neck following childhood irradiation. Prospective Studies. Medicine. 1985; 64:1-15.

28 Schneider, A. B.; Recant, W.; Pinskly, S. M., et al. Radiation-induced thyroid carcinoma. Ann Int Med. 1986; 105:405-412.

29 Shore, R. E.; Hildreth, N.; Dvoretsky, P; Pasternack, B., and Andrese, E. Benign thyroid adenomas among persons X-irradiated in infancy for enlarged thymus glands. Radiation Research. 1993 May; 134(2):217-23.

30 Ron, E.; Modan, B.; Pretson, D.; Alfandary, E.; Stovall, M., and Boice, Jr. J. D. Thyroid neoplasia following low-dose radiation in childhood. Radiation Research. 1989; 120:516-531.

31 Pottern, L. M.; Kaplan, M. M.; Larsen, P. R.; Silva, J. E.; Koenig, R. J.; Lubin, J. H.; Stovall, M., and Boice, J. D. Jr. Thyroid nodularity after childhood irradiation for lymphoid hyperplasia: a comparison of questionnaire and clinical findings. Journal of Clinical Epidemiology. 1990; 43

32 Schneider, A. B.; Ron, E.; Lubin, J.; Stovall, M., and Gierlowski, T. C. Dose-Response Relationships for Radiation-Induced Thyroid Cancer and Thyroid Nodules: Evidence for the Prolonged Effects of Radiation on the Thyroid. Journal of Clinical and Endocrinology and Metabolism. 1993; 77:362-369.

33 Prentice, R. L.; Kato, H.; Yoshimoto, K., and Mason, M. Radiation exposure and thyroid cancer incidence among Hiroshima and Nagasaki residents. Natl Cancer Inst Monogr. 1982; 62, 207-212

34 Thompson, D. E., et al. Cancer Incidence in Atomic Bomb Survivors. Part II: Solid tumors. Radiation Research. 1994; 137:s17-s67.

35 Friedlander, A. Status lymphaticus and enlargement of the thymus: With report of a case successfully treated by the x-ray. Arch Pediatr. 1907; 24:490-501.

36 Induction of thyroid cancer by ionizing radiation. NCRP Report 80. Washington, DC: Academy Press; 1985.

37 Hemplemann, L. H.; Hall, W. J.; Phillips, M., et al. Neoplasms in persons treated with x-rays in infancy: Fourth survey in 20 years. JNCI. 1975; 55:519-530.

38 Maxon, H. R.; Saenger, E. L.; Thomas, S. R., et al. Clinically important radiation-associated thyroid disease: A controlled study. JAMA. 1980; 244:1802-1807.

39 Shore, R. E.; Albert, R. E., and Pasternack, B. S. Follow-up study of patients treated by x-ray epilation for tinea capitis; re-survey for post-treatment illness and mortality experience. Arch Environ Health. 1976; 31:21.

40 Frohman, L. A.; Schneider, A. B.; Favus, M. J.; Stachura, M. E.; Arnold, J., and Arnold, M. "Thyroid carcinoma after head and neck irradiation. Evaluation of 1476 patients." (page 5), and, "Risk factors associated with the development of thyroid carcinoma and of nodular thyroid disease following head and neck irradiation." (page 231). IN Radiation-Associated Thyroid Carcinoma. DeGroot, L. J.; Grohman, L. A.; Kaplan, E. I., and Refetoff, S., Editors. Grune & Stratton, New York, New York, 1977.

41 De Groot, L. J.; Reilly, M.; Pinnameneni, K., et al. Retrospective and prospective study of radiation-induced thyroid disease. Am J Med. 1983; 74:852-862.

42 Ron, E; Lubin, J. H; Shore, R. E; Mabuchi, K; Modan, B; Potter, L. M; Schneider, A. B; Tucker, M. A, and Boice, J. D. Thyroid cancer after exposure to external radiation: A pooled analysis of seven studies. Radiation Research. 1995; 141:259-277.

43 Doniach, I. The effect of radioiodine alone and in combination with methylthiouracil upon tumor production in rat's thyroid gland. Br J Cancer. 1974; 30:487-495.

44 Doniach, I. Effects including carcinogenesis of I-131 and x-rays on the thyroid of experimental animals: a review. Health Phys. 1963; 9:1357-1362.

45 Lee, W; Chiacchierini, R. P; Shleine, B., and Telles. Thyroid tumors following I-131 or localized X irradiation to the thyroid and pituitary glands in rats. Radiation Research. 1982; 92:307-319.

46 Dobyns, B. M.; Sheline, G. E.; Workman, J. G.; Topkins, E. A.; McConahey, W. M., and Becker, D. V. Malignant and benign neoplasms of the thyroid in patients treated for hyperthyroidism: A report of the Cooperative Thyrotoxicosis Study. J Clinical Endocrinology. 1974; 38:976.

47 Holm, L.E.; Dahlquist, I.; Israelsson, A.; Lundell, G. Malignant thyroid tumors after I-131 therapy. New Engl J Med, 1980b, 303-188.

48 Holm, L.E. Malignant disease following iodine-131 therapy in Sweden. IN Radiation Carcinogenesis: Epidemiology and Biologic Significance. Boice, J.D., Jr., Fraumeni, J.F., Jr., Editors. Raven Press, New York, New York, 1984, page 263.

49 Safa, A.M.; Schumacker, O.P; and Rodrigues-Antunez, A. Long-term follow-up results in children and adolescents treated with radioactive iodine (I-131) for hyperthyroidism. New Engl J Med, 1975; 292:167.

50 Thompkins, E. Personal communications to HR Maxon, July 8 and August 26, 1976.

51 Ron, E., et al. Cancer Mortality following Treatment for Adult Hyperthyroidism. JAMA. 1998; 280:347-355.

52 Franklyn, JA, P Maisonneuve, M Sheppard, J Betteridge, P Boyle. Cancer incidence and mortality after radioiodine treatment for hyperthyroidism: a population-based cohort study [see comments]. Lancet 353[9170], 2111-5. 1999

53 Holm, L. E.; Lundell, G., and Wallinder, G. Incidence of malignant thyroid tumors in humans after exposure to diagnostic doses of iodine-131: I. Retrospective cohort study. JNCI. 1980a; 64:1055.

54 Holm, L. E.; Eklund, G., and Lundell, G. Incidence of malignant tumors in humans after exposure to diagnostic doses of iodine-131: II. Estimation of actual thyroid gland size, thyroidal radiation dose and predictive versus observed number of malignant thyroid tumors. JNCI. 1981; 64:1221.

55 Harris, B. S. H. Letter To: Maxon, H.R. 1980 Apr 30.

56 Hall, P; Mattson, A, and Boice, J. D. Thyroid cancer after diagnostic administration of iodine-131. Radiation Research. 1996; 145:86-92.

57 Hamilton, P. M; Chiacchierina. R.P, and Kaczmarek, R. G. "A follow-up study of persons who had iodine-131 and other diagnostic procedures during childhood and adolescence." Rockville, MD: US Dept HHS; 1989 Aug.

58 Rallison, M. L.; Dobyns, B. M.; Keating, F. R.; Rall, J. E., and Tyler, F. H. Thyroid diseases in children. A survey of subjects potentially exposed to fallout radiation. Am J Med. 1974; 56:457

59 Rallison, M.L.; Dobyns, B.M.; Keating, F.R.; Rall, J.E.; and Tyler, F.H. Thyroid nodularity in children. JAMA. 1975; 233:1069.

60 Kerber, R. A.; Till, J. E.; Simon, S. L.; Lyon, J. L.; Thomas, D. C.; Preston-Martin, S.; Rallison, M. L.; Lloyd, R. D., and Stevens, W. A cohort study of thyroid disease in relation to fallout from nuclear weapons testing. JAMA. 1993 Nov 3; 270(17):2076-82; ISSN: 0098-7484.

61 Conard, R. A. and et al. "Review of Medical Findings in a Marshallese Population Twenty-six Years After Accidental Exposure to Radioactive Fallout." Upton New York: Brookhaven National Laboratory; 1980; BNL 51261.

62 Conard, R. A. Late radiation effects in Marshall Islanders exposed to fallout 28 years ago. IN Radiation Carcinogenesis: Epidemiology and Biologic Significance. Boice, J. D. and Fraumeni, J. F. Jr., Editors. New York, New York: Raven Press; 1984; p. 57.

63 James, R. A. "Estimate of Radiation Dose to Thyroids of the Rongelap Children Following the BRAVO Event (Publication 12-273)." Livermore, University of California Radiation Laboratory: U.S. Department of Energy; 1964.

64 Lessard, E.; Miltenberger, R.; Conard, R., et al. "Thyroid Absorbed Dose for People at Rongelap, Utrik, and Sifo on March 1, 1954" (Publication BNL 51-882). Upton, New York, Brookhaven National Laboratory: U.S. Department of Energy; 1985.

65 Howard, J. E.; Vaswani, A., and Heotis, P. Thyroid disease among the Rongelap and Utirik population--an update. Health Physics. 1997 Jul; 73(1)190-8; ISSN: 0017-9078.

66 Simon, S. L. and Graham, J. C. Dose assessment activities in the Republic of the Marshall Islands. Health Physics. 1996 Oct; 71(4): 438-456; ISSN: 0017-9078.

67 Takahashi, T.; Trott, K. R.; Fujimori, K.; Simon, S. L.; Ohtomo, H.; Nakashima, N.; Takaya, K.; Kimura, N.; Satomi, S., and Schoemaker, M. J. An investigation into the prevalence of thyroid disease on Kwajalein Atoll, Marshall Islands. Health Physics. 1997 Jul; 73(1):199-213; ISSN: 0017-9078.

68 Harley, JH; Hallden, NA, and Ong, LD. "Summary of gummed film results, 1959." 15th ed.. New York: U.S. Atomic Energy Commission, New York Operations Office, Health and Safety Laboratory; 1960; HASL-93, UC 41, Health and Safety, TID-4500.

69 Simon, SL and Graham, JC. Findings of the first comprehensive radiological monitoring program of the Republic of the Marshall Islands. Health Physics. 1997; 73:66-87.

70 Hamilton, T. E.; van Belle, G., and LoGerfo, J. P. Thyroid neoplasia in Marshall Islanders exposed to nuclear fallout. JAMA. 1987 Aug 7; 258, (5): 629-636.

71 Baverstock, K.F. Thyroid cancer in children in Belarus after Chernobyl. World Health Stat Q. 993;46: 204-8.

72 Kazakov, V. S.; Demidchik, E. P., and Astakhova, L. N. Thyroid cancer after Chernobyl [letter]. Nature. 1992 Sep 3; 359(6390):21; ISSN: 0028-0836.

73 Pacini, F.; Vorontsova, T.; Demidchik, E. P.; Molinaro, E.; Agate, L.; Romei, C.; Shavrova, E.; Cherstvoy, E. D.; Ivashkevitch, Y.; Kuchinskaya, E.; Schlumberger, M.; Ronga, G.; Filesi, M., and Pinchera, A. Post-Chernobyl thyroid carcinoma in Belarus children and adolescents: comparison with naturally occurring thyroid carcinoma in Italy and Grance. Journal of Clinical Endocrinology & Metabolism. 1997 Nov; 82(11):3563-9; ISSN: 0021-972X.

74 Tronko, N.; Bogdanova, T.;Kommisarenko, I.; et al. "Thyroid cancer in children and adolescents in Ukraine after the Chernobyl accident (1986-1995)." IN The Radiological consequences of the Chernobyl Accident, Karaoglou, A.; Sobolev, BG.; Kairo, I.A., et al, Editors. Luxembourg ,1996, European Commission, ERU 16544 EN, pp. 683-690.

75 Likhtarev, AI; Sobolev, BG; Kairo, IA, et al. Thyroid cancer in the Ukraine. Nature. 1995 375:365.

76 Jacob, PGWF Heal. Thyroid cancer risk to children calculated. Nature 392[ ], 31-32. 1998

77 Sobolev, B.; Heidenreich, W. F.; Kairo, I.; Jacob, P.; Goulko, G., and Likhtarev, I. Thyroid cancer incidence in the Ukraine after the Chernobyl accident: comparison with spontaneous incidences. Radiation & Environmental Biophysics. 1997 Sep; 36(3):195-9; ISSN: 0301-634X.

78 National Academy of Sciences. Committee on an Assessment of Centers for Disease Control and Prevention for Disease Control and Prevention Radiation Studies from DOE Contractor Sites, Subcommittee to Review the Hanford Thyroid Disease Study Results and Final Report. National Academy Press, Washington, D.C.Â 1999.

79 Astakhova, Beal. Chernobyl-related thyroid cancer in children of Belarus: Âyroid cancer in children of Belarus: a case-control study . Radiation Research 150[ ], 349-356. 1998.

80 Stepanenko, V.; Gavrilin, Y.; Khorousch, V, et al. "The reconstruction of thyroid dose following Chernobyl." IN The radiological consequences of the Chernobyl accident, (Meeting dates 1996 Mar 18-1996 Mar 22 Minsk, Belarus); Karaoglou, A.; Desment, G.; Kelly, G. N., and Menzel, H. G., Editors. European Commission and the Belarus, Russian, and Ukranian Ministries on Chernobyl Affairs, Emergency Situations and Health; 1996.

81 Gavrilin, Y.; Drozdovitch, V.; Bouville, A.; Anspaugh, L., et al. "Estimation of Thyroid Doses Received by the Population of Belarus as a Results of the Chernobyl Accident" IN The Radiological Consequences of the Chernobyl Accident; (Meeting dates1996 Mar 18-1996 Mar 22. Minsk, Belarus); Karaoglou, A.; Desmet, G.; Kelly, G. N., and Menzel, H. G., Editors. European Commission and the Belarus, Russian, and Ukranian Ministries on Chernobyl Affairs, Emergency Situations, and Health; 1996.

82 Likhtarev, I.; Tabachny, L.; Jacob, P., et al. "Results of Large Scale Thyroid Dose Reconstruction in Ukraine". IN The radiological consequences of the Chernobyl accident; (Meeting dates 1996 Mar 18-1996 Mar 22, Minsk, Belarus.) Karaoglou, A.; Desmet, G.; Kelly, G. N., and Menzel, H. G., Editors. European Commision and the Belarus, Russian and Ukranian Ministries on Chernobyl Affairs, Emergency Situations and Health; 1996.

83 Health effects of exposure to low levels of ionizing radiation (BEIR V:1990). Washington, DC: National Academy Press; 1990; National Research Council.

84 Ron, E, J H Lubin, R E Shore, K Mabuchi, B Modan, L M Pottern, A B Schneider, M A Tucker, J D Boice Jr. Thyroid cancer after exposure to external radiation: a pooled analysis of seven studies. Radiation Research 141[3], 259-77. 1995

85 de Vathaire, F et al. Long-term effects on the thyroid of irradiation for skin angiomas in childhood. Radiation Research. 1993; 133:381-386.

86 Maxon, HR, S R Thomas, E L Saenger, C R Buncher, J G Kereiakes. Ionizing irradiation and the induction of clinically significant disease in the human thyroid gland. [Review] [59 refs]. American Journal of Medicine 63[6], 967-78. 1977.

87 Rallison, ML, B M Dobyns, F R Keating, J E Rall, F H Tyler. Thyroid nodularity in children. JAMA 233, 1069. 1975.

88 Rosen, I. B.; Strawbridge, H. G., and Bain, J. A case of hyperparathyroidism associated with radiation to the head and neck area. Cancer. 1975; 36:1111-1114.

89 Tissel, L-E.; Carlsson, S.; Fjalling, M., et al. Hyperparathyroidism subsequent to neck irradiation. Cancer. 1985; 56:1529-1533.

90 Cohen, J.; Gierlowski, T. C., and Schneider, A. B. A prospective study of hyperparathyroidism in individuals exposed to radiation in childhood. JAMA. 1990 Aug 1; 264, (5): 581-584.

91 Schneider, AB; Gierlowski, TC, et al. Dose-response relationships for radiation-induced hyperparathyroidism. Journal of Clinical Endocrinology and Metabolism. 1995; 80:254-257.

92 Fujiwara, S.; Esaki, H.; Sposto, R.; Akiba, S.; Neriishi, K.; Kodama, K, et al. "Hyperparathyroidism among Atomic Bomb Survivors in Hiroshima, 1986-1988." RERF; 1990; TR 8-90, 1-12.

93 Triggs, S. M. and Williams, E. D. Irradiation of the thyroid as a cause of parathyroid adenoma. Lancet. 1977; 1:593-594.

94 Lindsay, S.; Potter, G. D., and Chaikoff, I. L. Radioiodine-induced thyroid carcinomas in female rats. Induction by low doses of radiation. Arch Pathol. 1963; 75:20-24.

95 Bondeson, A., Bondeson L., and Thompson, N. W. Hyperparathyroidism after treatment with radioactive iodine: Not only a coincidence? Surgery. 1989; 106:1025-1027.

96 Doniach, I. and Williams, E. D. Biologic effects of radiation on the thyroid. IN Werner's The Thyroid. 5th ed. 1986; p. 437.

97 Tan, GH, H Gharib. Thyroid incidentalomas: management approaches to nonpalpable nodules discovered incidentally on thyroid imaging. [Review] [45 refs]. Annals of Internal Medicine 126[3], 226-31. 1997.

98 Hahn, K, P Schnell-Inderst, B Grosche, L E Holm. Thyroid cancer after diagnostic administration of iodine-131 in childhood. Radiation 156[1], 61-70. 2001.

99 Ezzat, S, D A Sarti, D R Cain, G D Braunstein. Thyroid incidentalomas. Prevalence by palpation and ultrasonography. Archives of Internal Medicine 154[16], 1838-40. 1994.

100 Brander, A, P Viikinkoski, J Tuuhea, L Voutilainen, L Kivisaari. Clinical versus ultrasound examination of the thyroid gland in common clinical practice. Journal of Clinical Ultrasound 20[1], 37-42. 1992

101 Tomimori, EK, R Y Camargo, H Bisi, G Medeiros-Neto. Combined ultrasonographic and cytological studies in the diagnosis of thyroid nodules. Biochimie 81[5], 447-52. 1999.

102 Woestyn, J, M Afshcrift, K Schelstraete, A Vermeulen. Demonstration of nodules in the normal thyroid by echography. Br J Radiol 58, 1179-82. 1985.

103 Tan, GH, H Gharib, C C Reading. Solitary thyroid nodule. Comparison between palpation and ultrasonography. Archives of Internal Medicine 155[22], 2418-23. 1995.

104 Brander, A, P Viikinkoski, J Nickels, L Kivisaari. Thyroid gland: us screening in a random adult population. Radiology 181[3], 683-7. 1991

105 Brander, AE, V P Viikinkoski, J I Nickels, L M Kivisaari. Importance of thyroid abnormalities detected at US screening: a 5-year follow-up. Radiology 215[3], 801-6. 2000.

106 Bruneton JN, Balu-Maestro C, Marcy PY, Melia P, Mourou MY. Very high frequency (13MHz) ultrasonographic examination of the normal neck. Journal of Ultrasound in Medicine 13[2], 87-90. 1994

107 Mortensen, M., Gross and Microscopic Findings in Clinically Normal Thyroid Glands. Unknown 15, 1270-1280. 1955.

108 Rago, T, P Vitti, L Chiovato, S Mazzeo, A De Liperi, P Miccoli, P Viacava, F Bogazzi, E Martino, A Pinchera. Role of conventional ultrasonography and color flow-doppler sonography in predicting malignancy in 'cold' thyroid nodules [see comments]. European Journal of Endocrinology 138[1], 41-6. 1998.

109 Takashima, S, H Fukuda, N Nomura, H Kishimoto, T Kim, T Kobayashi. Thyroid nodules: re-evaluation with ultrasound. Journal of Clinical Ultrasound 23[3], 179-84. 1995.

110 Sakaguchi, T, A Arakawa, M Takahashi. Appropriate use of ultrasonography in the neck. [Review] [35 refs]. Seminars in Roentgenology 35[1], 54-62. 2000.

111 Hegedus, L, S Karstrup. Ultrasonography in the evaluation of cold thyroid nodules [comment]. [Review] [9 refs]. European Journal of Endocrinology 138[1], 30-1. 1998.

112 Schneider, AB, C Bekerman, J Leland, J Rosengarten, H Hyun, B Collins, E Shore-Freedman, T C Gierlowski. Thyroid nodules in the follow-up of irradiated individuals: comparison of thyroid ultrasound with scanning and palpation. Journal of Clinical Endocrinology & Metabolism 82[12], 4020-7. 1997.

113 Antonelli, A, G Silvano, F Bianchi, C Gambuzza, L Tana, G Salvioni, V Baldi, L Gasperini, L Baschieri. Risk of thyroid nodules in subjects occupationally exposed to radiation: a cross sectional study [see comments]. Occupational & Environmental Medicine 52[8], 500-4. 1995.

114 Sugenoya, A, K Asanuma, Y Hama, H Masuda, G S Skidanenko, A T Anatoliebna, K Koike, A Komiyama, F Iida. Thyroid abnormalities among children in the contaminated area related to the Chernobyl accident. Thyroid 5[1], 29-33. 1995.

115 Healy, J. W. (HW-07317); Zuerner, L. V. (HW-08429); Roos, L. C. (HW-08430); Adley, F. E. (HW-09864); Singlevich, W (HW-09496-DEL, HW-09871, HW-11333-DEL, HW-16015, HW-17150, HW-17409, HW-17657, HW-17944, HW-18242, HW-18447, Hw-18741, HW-19046, HW-19332, HW-19625, HW-19896); Paas, H. J., (HW-25866, HW-26493, HW-17510, HW-727641, HW-28009, HW-29514, HW-30174, HW-30744) and Turner, L.D. (HW-3-5402, HW-3-5511, HW-3-5521).

116 Beck, D. M.; Darwin, R. F.; Erickson, A. R., and Eckert, R. L. "Milk Cow Feed Intake and Mild Production and Distribution Estimates for Phase I (DRAFT)." Report to the Technical Steering Panel of the Hanford Environmental Dose Reconstruction Project. Richland, WA: Battelle Pacific Northwest Laboratories; 1990.

117 Henckel, Daral (Owner of Twin Cities Creamery, Kennewick, WA). Interview with Michele Stenjham, Ph.D. Prossor, WA; 1988 Mar.

118 "Summary: Radiation Dose Estimates from Hanford Radioactive Material Releases to the Air and the Columbia River. Revision 1." The Technical Steering Panel of the Hanford Environmental Dose Reconstruction Project; 1994 Apr 21.

119 Oddie, T. H.; Fisher, D. A.; McConahey, W. M., et al. Iodine intake in the United States: A reassessment. J Clin Endocr. 1970; 30:659-665.

120 Book, S. A. "Age-related variation in thyroidal exposures from fission-produced radioiodines. Developmental Toxicology of Energy-Related Products." Washington, D.C.: U.S. Nuclear Regulatory Commission and Energy Research and Development Administration; 1978; Report No. Conf-771017.

121 Sikov, M. R. Effects of age on the iodine-131 metabolism and radiation sensitivity of the rat thyroid. Rad Res. 1969; 38:449.

122 Book, S. A.; McNeill, D. A., and Spangler, W. L. Age and its influence on effects of iodine-131 in guinea pig thyroid glands. Rad Res. 1980; 81:254.

123 Christov, K. Radiation-induced thyroid tumors in infant rats. Rad Res. 1978; 73:330.

124 Shore, R.E.;Woodward, E.D.; Hempleman, L.H. "Radiation-induced thyroid cancer." IN Radiation Carcinogenesis: Epidemiology and Biological Significance. New York, Raven Press, 1984.

125 Wang, JX, L A Zhang, B X Li, Y C Zhao, Z Q Qant, J Y a A T Zhang. Cancer Incidence and Risk Estimation Among Medical X-ray Workers in China, 1950-1995. Health Physics, 2002.

126 Napier B, EPRJHL. Responses to National Academy of Sciences Review Comments on Dosimetry in the Fred Hutchinson Cancer Research Center Comments on Dosimetry in the Fred Hutchinson Cancer Research Center Hanford Thyroid Disease Study. Battelle Pacific Northwest Laboratories, PNWD3060. [ ]. 2000.

127 Snyder, S, W Farris, B Napier, T Ikenberry, D Shipler, J Simpson. Parameters used in the environmental pathways and radiological dose modules (DESCARTES, CIDER, and CDR codes) of the Hanford Environmental Dose Reconstruction Integrated Codes (HEDRIC). 1994
.

128 Davis, Scott. Letter To: Barbara Aripa, et al. 1994 Jun 30.

129 Farris, WT; Napier, BA; Eslinger, PW, et al. "Atmospheric Pathway Dosimetry Report, 1944-1992." PNWD-2228 HEDR. 1994.

130 Farris, WT; Napier, BA, et al. Radiation doses from Hanford Site releases to the atmosphere and the Columbia River. Health Physics. 1996; 71(4):588-601.

131 Estimated Exposure and Thyroid Doses Received by the American People from Iodine-131 in Fallout Following Nevada Atmospheric Nuclear Bomb Tests. National Cancer Institute; 1997 Aug 1.

132 Ouderkirk, S. J. and Eslinger, P. W. Software Design Description for the Cider Dose Estimation Computer Code; Hanford Environmental Dose Reconstruction Project. Richland, Washington: Battelle Pacific Northwest Laboratory; 1993 Jul 8.

133 Stevens, W// Till, JE// Thomas, DC// Lyon JL, Kerber RA, Preston-Martin S, Simon SL, Rallison ML, Lloyd RD. Report of a Cohort Study of Thyroid Disease and Radioactive Fallout from the Nevada Test Site. 1992

134 Kerber, R. A. //Till, J. E. //Simon, S. L. //Lyon, J. L. //Thomas, D. C. //Preston-Martin, S. //Rallison, M. L. //Lloyd, R. D. //Stevens, W., Department of Family and Preventive Medicine, University of Utah, Salt Lake City 84132. A cohort study of thyroid disease in relation to fallout from nuclear weapons testing. JAMA. 1993 Nov 3; 270 (17): 2076-2082

135 Davis, Scott. Letter To: Jim Thomas. 1995 Feb 21.

136 Hempelmann, LH, W J Hall, M Phillips, R A Cooper, W R Ames. Neoplasms in persons treated with x-rays in infancy: fourth survey in 20 years. Journal of the National Cancer Institute 55[3], 519-30. 1975.

137 Shipler, DB; Napier, BA; Farris, WT, and Freshley, MD. Hanford Environmental Dose Reconstruction Project - An overview. *Health Physics*. 1996; 71(4):532-544.

138 National Research Council. Review of the Hanford Thyroid Disease Study Draft Final Report. 2000.

139 Mallick B. Semiparametric regression modeling with mixtures of Berkson and classical error, with application to fallout from the Nevada Test Site. Biometrics 58[1], 13-20. 2002.

140 Figge J., 2000, Thyroid Cancer. A Comprehensive Guide to Clinical Management, Wartofsky L. (ed), Humana Press.

141 National Council on Radiation Protection and Measurements. Induction of Thyroid Cancer by Ionizing Radiation. NCRP Report No. 80. Bethesda, MD. 1985.

142 Wang, C, L M Crapo. The epidemiology of thyroid disease and implications for screening. [Review] [142 refs]. Endocrinology & Metabolism Clinics of North America 26[1], 189-218. 1997.

143 Jorde, R, J Sundsfjord. Primary hyperparathyroidism detected in a health screening. Journal of Clinical Epidemiology 53[11], 1164-9. 2000.

144 Vander, JBGEA et al. Significance of Solitary Nontoxic Thyroid Nodules. New England Journal of Medicine 251[24], 970-973. 1954.

145 Vander, JB, E A Gaston, T R Dawber. The significance of nontoxic thyroid nodules. Final report of a 15-year study of the incidence of thyroid malignancy. Annals of Internal Medicine 69[3], 537-40. 1968.

146 Tunbridge, WM, D C Evered, R Hall, D Appleton, M Brewis, F Clark, J G Evans, E Young, T Bird, P A Smith. The spectrum of thyroid disease in a community: the Whickham survey. Clinical Endocrinology 7[6], 481-93. 1977.

147 Vanderpump, MP, W M Tunbridge, J M French, D Appleton, D Bates, F Clark, J Grimley Evans, D M Hasan, H Rodgers, F Tunbridge, a l et. The incidence of thyroid disorders in the community: a twenty-year follow-up of the whickham survey. Clinical Endocrinology 43[1], 55-68. 1995.

148 Schlesinger, Mea. Studies in Nodular Goiter. Incidence of thyroid nodules in routine necropsies in a nongoitrous region. Journal of the American Medical Association 110, 1638. 1938.

149 Oertel, Jea. Structural changes in the thyroid glands of healthy young men. The Medical Annals of the District of Columbia 34, 77-end. 1965

150 Rice, C. Incidence of nodules in the thyroid. Archives of surgery 24, 505-end. 1932

151 Hull, Oetal. Critical Analysis of Two Hundred Twenty-One Thyroid Glands. A.M.A. Archives of Pathology 59, 291-end. 1955.

152 Sawin, CT, W P Castelli, J M Hershman, P McNamara, P Bacharach. The aging thyroid. Thyroid deficiency in the Framingham Study. Archives of Internal Medicine 145[8], 1386-8. 1985.

153 Aghini-Lombardi, F, L Antonangeli, E Martino, P Vitti, D Maccherini, F Leoli, T Rago, L Grasso, R Valeriano, A Balestrieri, A Pinchera. The spectrum of thyroid disorders in an iodine-deficient community: the Pescopagano survey. Journal of Clinical Endocrinology & Metabolism 84[2], 561-6. 1999.

154 Andersen-Ranberg, K, B Jeune, M Hoier-Madsen, L Hegedus. Thyroid function, morphology and prevalence of thyroid disease in a population-based study of Danish centenarians. Journal of the American Geriatrics Society 47[10], 1238-43. 1999.

155 Canaris, GJ, N R Manowitz, G Mayor, E C Ridgway. The Colorado thyroid disease prevalence study. Archives of Internal Medicine 160[4], 526-34. 2000.

156 Hollowell, JBLSCAetal. Serum TSH, T4, and Thyroid Antibodies in the United States Population: Nhanes III. 105.

157 Spencer CA. Takeuchi M. Kazarosyan M. Wang CC. Guttler RB. Singer PA. Fatemi S. LoPresti JS. Nicoloff JT. Serum thyroglobulin autoantibodies: prevalence, influence on serum thyroglobulin measurement, and prognostic significance in patients with differentiated thyroid carcinoma. Journal of Clinical Endocrinology & Metabolism 83[4], 1121-7. 1998.

158 Hoffman DA. Radiation carcinogensis: Epidemiology and biological significance. 1984.

159 Prentice, LM, D I Phillips, D Sarsero, K Beever, S M McLachlan, B R Smith. Geographical distribution of subclinical autoimmune thyroid disease in britain: a study using highly sensitive direct assays for autoantibodies to thyroglobulin and thyroid peroxidase. Acta Endocrinologica 123[5], 493-8. 1990.

160 Lazarus, JH, M L Burr, A M McGregor, A P Weetman, M Ludgate, J S Woodhead, R Hall. The prevalence and progression of autoimmune thyroid disease in the elderly. Acta Endocrinological 106[2], 199-202. 1984.

161 Ericsson, UB, F Lindgarde. The epidemiology of thyroid disorders in a seaport community in southern Sweden. Journal of Clinical Epidemiology 43[7], 645-50. 1990.
162 Christensson, T, K Hellstrom, B Wengle, A Alveryd, B Wikland. Prevalence of hypercalcaemia in a health screening in Stockholm. Acta Medica Scandinavica 200[1-2], 131-137. 1976.

163 Duquenne, M, V Rohmer, S Guyetant, G Becouarn, N W Barbot, J P Saintandre, J Ronceray, J C Bigorgne. Solitary thyroid nodule, a comparative study of cytology, scintigraphy and histologic findings [french]. Presse Medicale 1997 Apr 5[26], 11.

164 Fujiwara, S, R Sposto, H Ezaki, S Akiba, K Neriishi, K Kodama, Y Hosoda, K Shimaoka. Hyperparathyroidism among atomic bomb survivors in Hiroshima. Radiation Research 130[3], 372-8. 1992.

165 Lundgren, E, J Rastad, E Thrufjell, G Akerstrom, S Ljunghall. Population-based screening for primary hyperparathyroidism with serum calcium and parathyroid hormone values in menopausal women. Surgery 121[3], 287-94. 1997.

166 Jorde, R, Naa, K H , J Sundsfjord. Primary hyperparathyroidism detected in a health screening. Journal of Clinical Epidemiology 53[11], 1164-9. 2000.